Title
The ζ(2) limit in the random assignment problem

Journal
Random Structures and Algorithms, 18(4)

ISSN
1042-9832

Author
Aldous, David J

Publication Date
2001-07-01

DOI
10.1002/rsa.1015

Peer reviewed

The $\zeta(2)$ Limit in the Random Assignment Problem

David J. Aldous*
Department of Statistics
University of California
367 Evans Hall # 3860
Berkeley CA 94720-3860
aldous@stat.berkeley.edu
http://www.stat.berkeley.edu/users/aldous

February 27, 2001

Abstract

The random assignment (or bipartite matching) problem asks about $A_n = \min_\pi \sum_{i=1}^n c(i, \pi(i))$, where $(c(i,j))$ is a $n \times n$ matrix with i.i.d. entries, say with exponential(1) distribution, and the minimum is over permutations π. Mézard and Parisi (1987) used the replica method from statistical physics to argue non-rigorously that $EA_n \to \zeta(2) = \pi^2/6$. Aldous (1992) identified the limit in terms of a matching problem on a limit infinite tree. Here we construct the optimal matching on the infinite tree. This yields a rigorous proof of the $\zeta(2)$ limit and of the conjectured limit distribution of edge-costs and their rank-orders in the optimal matching. It also yields the asymptotic essential uniqueness property: every almost-optimal matching coincides with the optimal matching except on a small proportion of edges.

Key words and phrases. Assignment problem, bipartite matching, cavity method, combinatorial optimization, distributional identity, infinite tree, probabilistic analysis of algorithms, probabilistic combinatorics, random matrix, replica method, replica symmetry.

AMS subject classifications. 05C80, 60C05, 68W40, 82B44.

*Research supported by N.S.F. Grant DMS99-70901

1 Introduction

1.1 Background

> Consider the task of choosing an assignment of n jobs to n machines in order to minimize the total cost of performing the n jobs. The basic input for the problem is an $n \times n$ matrix $(c(i,j))$, where $c(i,j)$ is viewed as the cost of performing job i on machine j, and the *assignment problem* is to determine a permutation π that solves
>
> $$A_n = \min_{\pi} \sum_{i=1}^{n} c(i, \pi(i)).$$
>
> No doubt the simplest stochastic model for the assignment problem is given by considering the $c(i,j)$ to be independent random variables with the uniform distribution on $[0,1]$. This model is apparently quite simple, but after some analysis one finds that it possesses a considerable richness. *Steele [31].*

Part of this richness is that there have been four different approaches to the study of the random assignment problem, which we shall describe briefly. See [7, 31] for further history. Our focus is upon mathematical properties of A_n and the minimizing π, rather than algorithmic questions.

(a) *Rigorous bounds via linear programming.* Steele ([31] Chapter 4) goes on to give a detailed account of results of Walkup [35], Dyer - Frieze - McDiarmid [9] and Karp [14] which lead to the upper bound $EA_n \leq 2$. The lower bound $\limsup_n EA_n \geq 1 + e^{-1}$ was proved by Lazarus [18] and subsequent work of Olin [25] and Goemans - Kodialam [11] improved the lower bound to 1.51. Recently Coppersmith - Sorkin [7] improved the upper bound to 1.94.

(b) *The replica method.* Mézard - Parisi [23] gave a non-rigorous argument based on the *replica method* [24] to show $EA_n \to \pi^2/6$. The replica method and the somewhat related *cavity method* is extensively used in statistical physics in non-rigorous analysis of spin glass and related disordered systems [26], and there is considerable interest in mathematical formalization of the method (see e.g. Talagrand [32]).

(c) *Conjectured exact formulas.* Parisi [27] conjectured that in the case where the $c(i,j)$ have exponential(1) distribution, there is the exact formula

$$EA_n = \sum_{i=1}^{n} i^{-2}. \tag{1}$$

It is natural to seek an inductive proof. Coppersmith - Sorkin [7] and others [4, 6, 19] have recently formulated, and verified for small n, more general conjectures concerning exact formulas for related problems, e.g. involving incomplete matches on $m \times n$ bipartite graphs. But (1) has not been proved rigorously (an argument of Dotsenko [8] is incomplete).

(d) *Weak convergence.* Aldous [2] proved that $\lim_n EA_n$ exists, by identifying the limit as the value of a minimum-cost matching problem on a certain random weighted infinite tree.

1.2 Results

In this paper we continue the "weak convergence" analysis mentioned above. By studying the infinite-tree matching problem, we obtain a rigorous proof of the Mézard - Parisi conclusion $EA_n \to \pi^2/6$ (Theorem 1), and also of their formula for the limit distribution of edge-costs in the optimal assignment (Theorem 2), and of a conjecture of Houdayer et al [12] concerning order-ranks of edges in the optimal assignment (Theorem 3). Theorem 4 introduces and proves an *asymptotic essential uniqueness* (AEU) property of the random assignment problem: roughly, AEU asserts that every almost-optimal matching coincides with the optimal matching except on a small proportion of edges. Studying AEU is interesting for two reasons. First, it can be defined for many "optimization over random data" problems (section 7.1), providing a theoretical classification of such problems (AEU either holds or fails in each problem) somewhat in the spirit of computational complexity theory. Second, in the statistical physics of disordered systems it has been suggested [24] that the minima of the Hamiltonian should typically have an ultrametric structure, suggesting that in the associated optimization problem the AEU property should fail. Theorem 4 thus shows that the random assignment problem (studied by physicists as a toy model of disordered systems) has qualitatively different behavior than that predicted for more realistic models.

It follows easily from [2] that the distribution of the $c(i,j)$ affects the limit of EA_n only via the value of the density function of $c(i,j)$ at 0 (assuming this exists and is strictly positive: see section 6.1 for the power-law case), and so in particular we may assume that the distribution is exponential(1) instead of uniform. Indeed, to avoid later normalizations it is convenient to start out by rephrasing the problem as

$$A_n = \min_\pi \tfrac{1}{n} \sum_{i=1}^n c(i, \pi(i))$$

where the $(c(i,j))$ are independent with exponential distribution with mean n. Write π_n for the permutation attaining the minimum.

Theorem 1 *$\lim_n EA_n = \pi^2/6$.*

Theorem 2 *$c(1,\pi_n(1))$ converges in distribution; the limit distribution has density*

$$h(x) = \frac{e^{-x}(e^{-x}-1+x)}{(1-e^{-x})^2},\ 0 \le x < \infty.$$

Theorem 3 *For each $k \ge 1$ define*

$$q_n(k) = P(c(1,\pi_n(1)) \textit{ is the k'th smallest of } \{c(1,1), c(1,2), \ldots, c(1,n)\}).$$

Then $\lim_n q_n(k) = 2^{-k}$.

Theorem 4 *For each $0 < \delta < 1$ there exists $\varepsilon(\delta) > 0$ such that, if μ_n are permutations (depending on $(c(i,j))$) such that $En^{-1}\#\{i : \mu_n(i) \ne \pi_n(i)\} \ge \delta$, then*

$$\liminf_n E\left(\frac{1}{n}\sum_{i=1}^{n} c(i,\mu_n(i))\right) \ge \pi^2/6 + \varepsilon(\delta).$$

The logical structure of our proof is as follows.

- Review from [2] the infinite-tree minimum-weight matching problem and its connection with the finite-n problem (sections 3 and 4.1).
- Describe one matching $\mathcal{M}_{\rm opt}$ in the infinite-tree setting (section 4.3).
- Calculate the cost of the matching $\mathcal{M}_{\rm opt}$ and give an expression for the extra cost of any other matching $\mathcal{M}$ (section 4.5).

In principle our method should yield an explicit lower bound for $\varepsilon(\delta)$ in Theorem 4 – see section 5.4.

Matchings on the infinite tree are required to satisfy a certain *spatial invariance* property, and this is the only technically difficult ingredient of the paper. Granted the abstract structure, the actual calculations underlying the formulas in Theorems 1 - 3 are quite straightforward, and we set out these calculations first in section 2. In a sense the calculation of the limit in Theorem 1 can be summarized in one sentence. Let X, X_i be i.i.d. with distribution determined by the identity

$$X \stackrel{d}{=} \min_i(\xi_i - X_i) \text{ where } (\xi_i) \text{ is a Poisson (rate 1) process on } [0,\infty); \quad (2)$$

then

$$\lim_n EA_n = \int_0^\infty x P(X_1 + X_2 > x) \, dx. \tag{3}$$

From these it is straightforward (section 2) to evaluate the limit as $\pi^2/6$, though we do not have any calculation-free explanation of why such a simple number should appear here or in Theorem 3.

We remark that the technically most difficult, and conceptually most important, part of the proofs of Theorems 1 - 4 is the result from [2] (restated as Theorem 11) connecting the finite-n assignment problem with the infinite-tree problem. The new additions of this paper are precisely the construction (section 4.3) of the matching $\mathcal{M}_{\mathrm{opt}}$ on the infinite tree, and its analysis in sections 4.4 and 4.5, together with the elementary calculations in section 2.

1.3 Discussion

This paper was of course motivated by a desire to make the Mézard - Parisi [23] result rigorous. In retrospect our method, based on weak convergence to a limit random structure, seems roughly similar to the non-rigorous *cavity method* from statistical physics (see [28, 32] for brief descriptions, and Talagrand [34] for its rigorous application in the Sherrington-Kirkpatrick model) though I do not understand that method well enough to judge the degree of similarity. So can our method be used for other mean-field models of disordered systems in which the cavity method has been used? It turns out it is easy to use our method to write down heuristic solutions to three variations of the random assignment problem:

- The case where the density function $f_c(\cdot)$ of the costs $c(i,j)$ has $f_c(x) \sim x^r$ as $x \downarrow 0$ (section 6.1).
- The one-parameter Gibbs measure associated with the assignment problem, recently studied rigorously by Talagrand [33] (section 6.3).
- The random combinatorial traveling salesman problem (section 6.2).

We shall compare heuristic solutions from our method with those from the cavity method. Making a rigorous proof in these variations involves precisely the issue in [2]: one needs a rigorous argument identifying the optimal solution of an infinite-tree problem as the limit of optimal solutions of finite-n problems.

It is worth noting that in our method, calculations of limit quantities depend on solving a problem-specific distributional identity (2, 39, 42, 46). Such identities arise frequently in probabilistic analysis of recursive algorithms [30] and "probability on trees" [20], and seem worthy of further systematic study.

Section 7 contains miscellaneous discussion; the AEU property in more general contexts, the (lack of) insight our results cast on the exact conjecture (1) and the intriguing analogy with Frieze's $\zeta(3)$ result.

Our method rests upon a kind of "local convergence" (13) of π_n to $\mathcal{M}_{\rm opt}$, saying that the structure of π_n and edge-costs within a finite distance (interpreting edge-costs as distances) of a typical vertex converges in distribution to the structure of $\mathcal{M}_{\rm opt}$ and edge-costs in the infinite tree. So in a sense it is Theorem 2 which is our basic result. We can deduce Theorem 1 because $EA_n = Ec(1, \pi_m(1))$, but we cannot study the variance of A_n by our methods. On the other hand we could write down expressions for the limit behavior of arbitrarily complicated "finite-distance" variations on Theorems 2 and 3, e.g. for

$$\begin{aligned}\gamma(x,y,z) &:= \lim_n \\ &\tfrac{1}{n} E\,(\text{ number of pairs } (i,k):\ c(i,\pi(i)) \le x, c(k,\pi(k)) \le y, c(i,\pi(k)) \le z)\,.\end{aligned}$$

Moreover, provided one could verify appropriate uniform integrability of higher moments conditions analogous to the first moment conditions in [2] (we haven't tried), one could deduce

$$\frac{1}{n}\sum_{i=1}^{n} c^r(i,\pi_n(i)) \to \int_0^\infty x^r h(x)\ dx, \quad 0 < r < \infty$$

where the limit has the intriguing exact formula (11).

Studying $n \to \infty$ asymptotics via an infinite tree approximation has, implicitly at least, a long history, in that many elementary results for the sparse random graph $\mathcal{G}(n, \alpha/n)$ are consequences of a local approximation by a Galton-Watson tree with Poisson(α) offspring. Karp and Sipser [15] treat matchings in that context.

2 Calculations

Consider the symmetric probability density

$$f_X(x) = \left(e^{x/2} + e^{-x/2}\right)^{-2}, \quad -\infty < x < \infty$$

which statisticians call the *logistic distribution.* It is standard [16] that the corresponding distribution function and variance are

$$F_X(x) := \int_{-\infty}^{x} f_X(y)dy = (1+e^{-x})^{-1}, \quad -\infty < x < \infty \tag{4}$$

$$\mathrm{var}\ X := \int_{-\infty}^{\infty} x^2 f_X(x)dx = \pi^2/3 \tag{5}$$

and that the logistic distribution is characterized by the property

$$f(x) = F(x)(1-F(x)), \quad -\infty < x < \infty. \tag{6}$$

In the calculations below we seek to exploit symmetry and structure, rather than rely on "brute force calculus".

Lemma 5 *Let $0 < \xi_1 < \xi_2 < \ldots$ be the points of a Poisson process of rate 1. Let $(X; X_i, i \geq 1)$ be independent random variables with some common distribution μ. Then*

$$\min_{1 \leq i < \infty} (\xi_i - X_i) \ \stackrel{d}{=}\ X \tag{7}$$

if and only if μ is the logistic distribution.

Proof. The points $\{(\xi_i, X_i)\}$ form a Poisson point process $\mathcal{P}$ on $(0,\infty) \times (-\infty,\infty)$ with mean intensity $\rho(z,x)dzdx = dz\ \mu(dx)$. [This is a continuous analog of the elementary *splitting* property: if the points of a Poisson (rate 1) process are marked by $r \in R$ independently with probability p_r then the resulting subprocesses are independent Poisson processes of rates p_r.] The distribution function F of μ satisfies

$$\begin{aligned}
1 - F(y) &= P\left(\min_{1\leq i<\infty}(\xi_i - X_i) \geq y\right) \text{ by (7)} \\
&= P\left(\text{no points of } \mathcal{P} \text{ in } \{(z,x) :\ z - x \leq y\}\right) \\
&= \exp\left(-\int\int_{z-x\leq y} \rho(z,x)\ dzdx\right) \\
&= \exp\left(-\int_0^\infty \bar{F}(z-y)\ dz\right), \text{ where } \bar{F}(y) = 1 - F(y) \\
&= \exp\left(-\int_{-y}^\infty \bar{F}(u)du\right).
\end{aligned}$$

Differentiating, we see this is equivalent to

$$F'(y) = \bar{F}(-y)\bar{F}(y). \tag{8}$$

This implies the density $F'(\cdot)$ is symmetric, and so (8) is equivalent to the condition (6) characterizing the logistic distribution. $\square$

In some subsequent calculations we use the general identity (for arbitrary real-valued r.v.'s V, W)

$$E(V - W)^+ = \int_{-\infty}^{\infty} P(V > x > W)\ dx \tag{9}$$

which follows from the fact $EU = \int_0^\infty P(U > u)\ du$ when $U \geq 0$ a.s.. Lemmas 6, 7 and 9 are the "calculation" parts of Theorems 1, 2 and 3 respectively.

Lemma 6 *Let X_1 and X_2 be independent random variables with the logistic distribution. Then*

$$h(x) := P(X_1 + X_2 > x), \quad 0 \leq x < \infty$$

is the density of a probability distribution on $[0, \infty)$ with mean $\pi^2/6$.

Proof. To show it is a *probability* density,

$$\begin{aligned}
\int_0^\infty h(x)dx &= E(X_1 + X_2)^+ \\
&= E(X_1 - X_2)^+ \text{ by symmetry} \\
&= \int_{-\infty}^{\infty} P(X_1 \geq y \geq X_2)\ dy \text{ by (9)} \\
&= \int_{-\infty}^{\infty} (1 - F(y))F(y)\ dy \\
&= \int_{-\infty}^{\infty} f(y)\ dy \text{ by (6)} \\
&= 1.
\end{aligned}$$

And the mean is

$$\begin{aligned}
\int_0^\infty xh(x)dx &= \int_0^\infty xP(X_1 + X_2 \geq x)\ dx \\
&= \tfrac{1}{2}E((X_1 + X_2)^+)^2 \\
&= \tfrac{1}{4}E(X_1 + X_2)^2 \text{ by symmetry} \\
&= \tfrac{1}{2}EX_1^2 \\
&= \pi^2/6 \text{ using (5)}.
\end{aligned}$$

□

The next lemma derives an explicit formula for $h(x)$, though it is not needed except to state the conclusion of Theorem 2 (my thanks to Boris Pittel for this calculation).

Lemma 7
$$h(x) = \frac{e^{-x}(e^{-x} - 1 + x)}{(1 - e^{-x})^2}; \quad 0 \leq x < \infty. \tag{10}$$

Proof.

$$\begin{aligned}
h(x) &= \int_{-\infty}^{\infty} \frac{1}{(e^{u/2} + e^{-u/2})^2} \, \frac{1}{1 + e^{x-u}} \, du \\
&= \int_{-\infty}^{\infty} \frac{e^u}{(e^u + 1)^2} \, \frac{e^u}{e^x + e^u} \, du \\
&= \int_0^{\infty} \frac{t}{(t+1)^2 (t + e^x)} \, dt \\
&= \int_0^{\infty} \left[\frac{e^x}{(e^x - 1)^2} \left(\frac{1}{t+1} - \frac{1}{t + e^x} \right) - \frac{1}{e^x - 1} \frac{1}{(t+1)^2} \right] dt \\
&= \frac{e^x}{(e^x - 1)^2} \left(\log \frac{t+1}{t + e^x} \right) \Bigg|_0^{\infty} + \frac{1}{e^x - 1} \, \frac{1}{t+1} \Bigg|_0^{\infty} \\
&= \frac{x e^x}{(e^x - 1)^2} - \frac{1}{e^x - 1}
\end{aligned}$$

and this equals the formula at (10). □

Remarks. (a) Mézard - Parisi ([21] eq. (33)) write down the formula

$$\tilde{h}(x) = \frac{x - e^{-x} \sinh x}{\sinh^2 x}$$

as the limit density for edge-costs in the optimal assignment, in the essentially equivalent non-bipartite matching problem on $2n$ vertices. One can check $\tilde{h}(x) = 2h(2x)$. The factor of 2 merely reflects the different normalization convention (dividing by $2n$ instead of n).

(b) As a corollary to extensive study of infinitely divisible laws associated with Lévy processes, Pitman and Yor ([29] eq. (131)) extend Lemma 6 to the following intriguing formula for all moments:

$$\int_0^{\infty} x^p h(x) \, dx = p\Gamma(p+1)\zeta(p+1), \quad 0 < p < \infty. \tag{11}$$

Lemma 8 *Let (X, X_1, X_2, η) be independent, the X's having logistic distribution and η having exponential(1) distribution. Then*

$$X \stackrel{d}{=} \min(X_1, X_2) + \eta. \tag{12}$$

Proof. In the setting of Lemma 5,

$$(\xi_1, \xi_2, \xi_3, \ldots) \stackrel{d}{=} (\eta, \eta + \xi'_1, \eta + \xi'_2, \ldots)$$

where the (ξ'_i) are a Poisson (rate 1) process independent of η. So identity (7) implies

$$X \stackrel{d}{=} \eta + \min(-X_1, \min_{i \geq 1}(\xi'_i - X_{i+1})).$$

But by (7) $\min_{i \geq 1}(\xi'_i - X_{i+1}) \stackrel{d}{=} X_2$, so

$$X \stackrel{d}{=} \eta + \min(-X_1, X_2)$$

and the result follows from the symmetry of the logistic distribution. □

Note that (12) does not characterize the logistic distribution, because $X+$ constant will also satisfy (12). But it is not hard to show (Antar Bandyopadhyay, personal communication) these are the only solutions.

Lemma 9 *In the setting of Lemma 5, for each $k \geq 1$*

$$\int_0^\infty P\left(x - X < \min_{1 \leq i < \infty}(\xi_i - X_i) \text{ and } x \geq \xi_k\right) \, dx = 2^{-k}.$$

Proof. Using symmetry of the logistic distribution, we may rewrite the integrand as $P(\xi_k \leq x < \min_{1 \leq i < \infty}(\xi_i + X_i) - X)$. Then by (9) the value ($Q(k)$, say) of the integral is

$$Q(k) = E\left(\min_{1 \leq i < \infty}(\xi_i + X_i) - (\xi_k + X)\right)^+.$$

Write

$$(\xi_1, \xi_2, \ldots) = (\xi_k - \eta_{k-1}, \xi_k - \eta_{k-2}, \ldots, \xi_k - \eta_1, \xi_k, \xi_k + \xi'_1, \xi_k + \xi'_2, \ldots),$$

so that (ξ'_i) is a Poisson process. Then

$$Q(k) = E\left(\min(X_1 - \eta_{k-1}, \ldots, X_{k-1} - \eta_1, X_k, \min_{1 \leq i < \infty}(\xi'_i + X_{k+i})) - X\right)^+$$

$$
\begin{aligned}
&= E(\min(X_1-\eta_{k-1},\ldots,X_{k-1}-\eta_1,X_k,X_{k+1}) \;-X)^+ \\
&\quad\text{by (7) and symmetry} \\
&= \int_{-\infty}^{\infty} P(X<x<\min(X_1-\eta_{k-1},\ldots,X_{k-1}-\eta_1,X_k,X_{k+1}))dx \text{ by (9)} \\
&= \int_{-\infty}^{\infty} F(x)P(x<\min(X_1-\eta_{k-1},\ldots,X_{k-1}-\eta_1,X_k))\;(1-F(x))dx \\
&= \int_{-\infty}^{\infty} f(x)P(x<\min(X_1-\eta_{k-1},\ldots,X_{k-1}-\eta_1,X_k))\;dx \text{ by (6)} \\
&= P(X<\min(X_1-\eta_{k-1},\ldots,X_{k-1}-\eta_1,X_k)).
\end{aligned}
$$

Since $(\eta_1,\ldots\eta_{k-1})$ are distributed as the first $k-1$ points of a Poisson process, we can relabel variables to obtain

$$Q(k)=P(X<\min(X_0,X_1-\xi_1,X_2-\xi_2,\ldots,X_{k-1}-\xi_{k-1})).$$

For $k=1$ this says $Q(1)=P(X<X_0)=1/2$, so it is enough to show $Q(k)=\frac{1}{2}Q(k-1)$ for $k\geq 2$. Write $M=\min(X,X_0)$. Then $P(X=M)=1/2$ and this event is independent of the value of M, so

$$Q(k)=\tfrac{1}{2}P(M<\min(X_1-\xi_1,X_2-\xi_2,\ldots,X_{k-1}-\xi_{k-1})).$$

Writing

$$(\xi_1,\xi_2,\xi_3,\ldots)\;\stackrel{d}{=}\;(\eta,\eta+\xi_1',\eta+\xi_2',\ldots)$$

and setting $X_i'=X_{i+1}$ gives

$$Q(k)=\tfrac{1}{2}P(M+\eta<\min(X_0',X_1'-\xi_1',\ldots,X_{k-2}'-\xi_{k-2}')).$$

But Lemma 8 shows $M+\eta\;\stackrel{d}{=}\;X$ and so we have proved $Q(k)=\frac{1}{2}Q(k-1)$.

3 The key picture

A central idea of our arguments is an *unfolding map*, which we describe informally here with the aid of figure 1, before starting the rigorous development in section 4.

Consider the complete bipartite graph G_{nn} on $n+n$ vertices, and the infinite ordered n-ary tree $\mathbf{T}^{(n)}$ with root ϕ and with other vertices labeled in the natural way as n-ary strings. Now suppose we are given edge-weights on G_{nn}, the weights being n^2 distinct positive real numbers. There are many ways to define a graph homomorphism from $\mathbf{T}^{(n)}$ onto G_{nn}. We use a particular one, the *folding map*, which pays attention to orderings of edge-weights. The inverse "unfolding map" then gives weights to the edges of $\mathbf{T}^{(n)}$, each weight being one of the n^2 weights on G_{nn}, and identifies each vertex of G_{nn} with an infinite subset of vertices of $\mathbf{T}^{(n)}$.

The top of figure 1 shows a realization of edge-weights on G_{nn}, for $n = 4$, and the bottom shows part of $\mathbf{T}^{(n)}$. First, an arbitrary vertex of G_{nn} is identified with ϕ. Then the vertices in G_{nn} at the other end of edges from ϕ are identified with vertices $1, 2, \ldots, n$ of $\mathbf{T}^{(n)}$, ordered so that the weights of edges (ϕ, i), $1 \leq i \leq n$ are increasing. Then for each i the $n-1$ vertices at the other end of edges from vertex i, where we exclude vertex ϕ, are identified with vertices $i1, i2, \ldots, i(n-1)$ of $\mathbf{T}^{(n)}$, ordered so that the weights of edges (i, ij), $1 \leq j \leq n-1$ are increasing. Continue: once a vertex of G_{nn} has been identified with a vertex v of $\mathbf{T}^{(n)}$, the children $v1, v2, \ldots, v(n-1)$ in $\mathbf{T}^{(n)}$ are identified with the vertices in G_{nn} linked by an edge to v, except for the vertex previously identified with the parent of v. In figure 1 we see that a particular vertex of G_{44} has been identified with the depth-2 vertices $11, 23, 32, 42$ of $\mathbf{T}_4$. The vertex identified with 1 will also be identified with the depth-3 vertices $231, 321, 421, 222, 312, 433, 212, 333, 412$.

Now the point of this construction is to obtain a structure within which it makes sense to let $n \to \infty$. The tree $\mathbf{T}^{(n)}$ is a subtree of the infinite-degree tree $\mathbf{T}$ described below, and the random edge-weights in the random assignment problem specify random weights on a subset of edges of $\mathbf{T}$. As $n \to \infty$ the distribution of edge-weights on $\mathbf{T}$ converge to a simple limit distribution of edge-weights with much independence (and no repeated values).

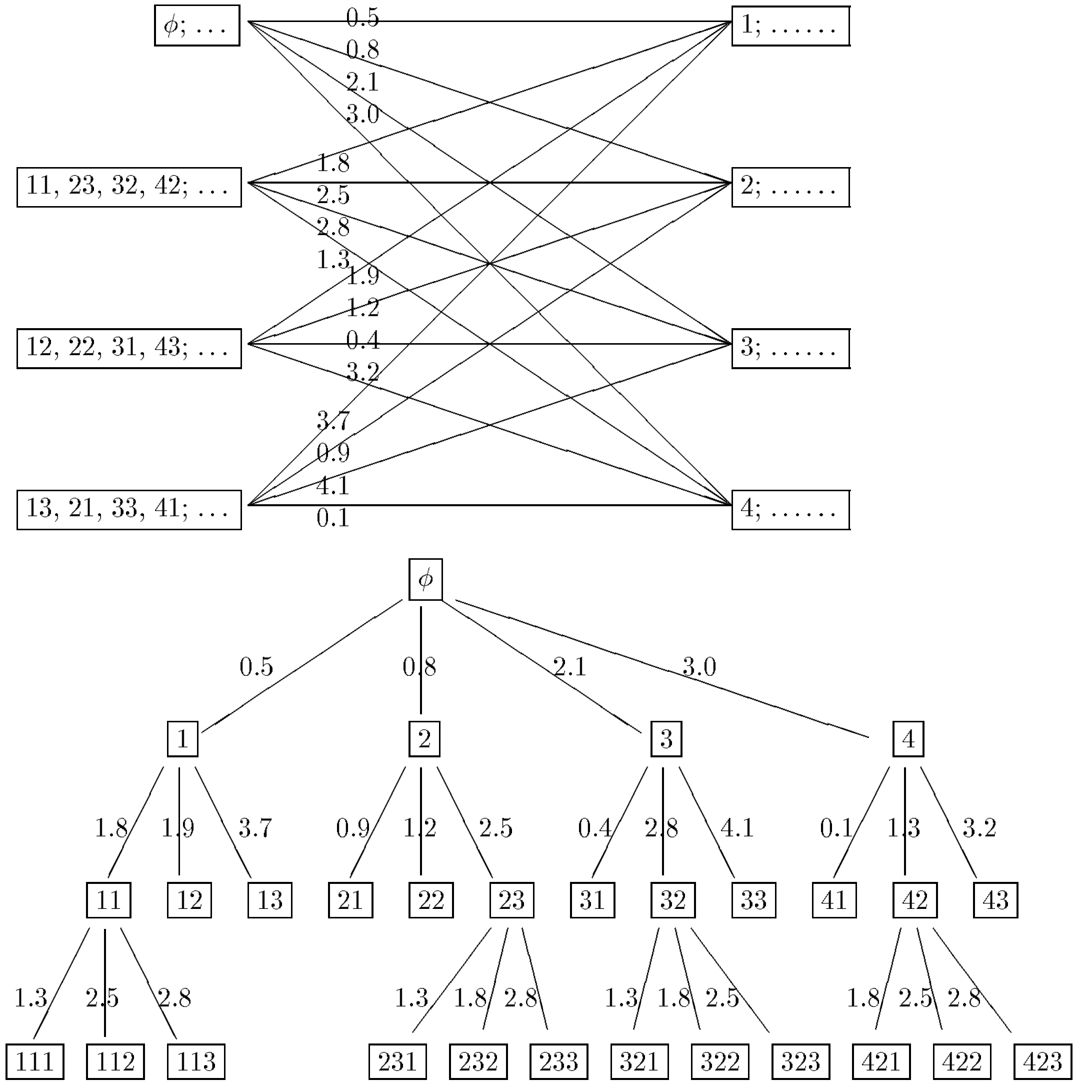

Figure 1. The unfolding map.

4 Working on the infinite tree

4.1 The infinite matching problem

In this section we review results from Aldous [2], which unfortunately require considerable space to describe.

Let $\mathbf{V}$ be the set of finite words $v = v_1 v_2 \ldots v_d$, $0 \leq d < \infty$ where each v_i is a natural number; include in $\mathbf{V}$ the empty word ϕ. There is a natural tree $\mathbf{T}$ with vertex-set $\mathbf{V}$ and edge-set $\mathbf{E}$, where an edge $e \in \mathbf{E}$ is of the form $e = (v, vj), j \geq 1$, where for $v = v_1 v_2 \ldots v_d$ we write $vj = v_1 v_2 \ldots v_d j$ for the j'th *child* of v, and call v the *parent* of vj. Now attach random edge-weights as follows. For each $v \in \mathbf{V}$ let the weights $(W(v, vj), j \geq 1)$ on the edges $((v, vj), j \geq 1)$ be the points of a Poisson (rate 1) point process on $(0, \infty)$, independently as v varies. Call this structure the *Poisson-weighted infinite tree* (PWIT). See figure 2 for an illustration. Write λ for the probability distribution of the whole configuration $(W(e))$ of edge-weights. So λ is a probability measure on the space $\mathbf{W} = (0, \infty)^{\mathbf{E}}$ of all possible configurations $\mathbf{w} = (w(e), e \in \mathbf{E})$ of edge-weights.

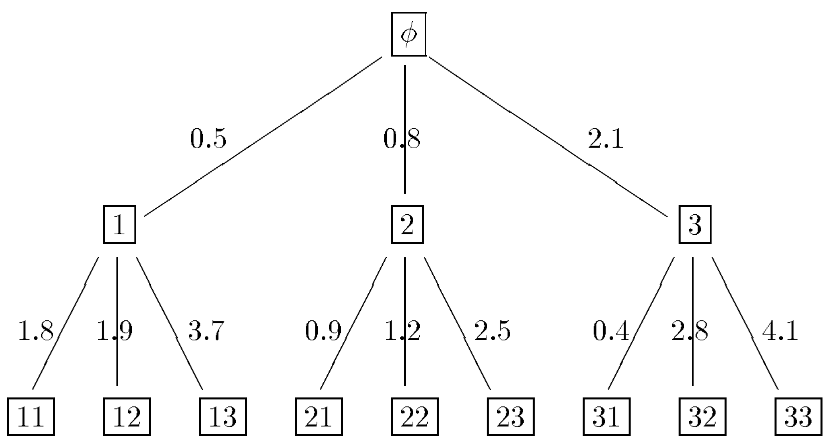

Figure 2. Part of a realization $\mathbf{w}$ of the PWIT. The weight $w(e)$ is written next to the edge e.

A *matching* $\mathbf{m}$ on the PWIT is a set of edges of $\mathbf{T}$ such that each vertex is incident to exactly one edge in the set. Formally we can identify the set $\mathbf{M}$ of matchings as the subset $\mathbf{M} \subseteq \{0,1\}^{\mathbf{E}}$ defined by: $\mathbf{m} = (m(e)) \in \mathbf{M}$ iff $\sum_{e: v \in e} m(e) = 1 \ \forall v \in \mathbf{V}$. A matching can also be regarded as a map from vertices to vertices, in which case we write $\vec{\mathbf{m}}(v) = v'$ to indicate (v, v') is an edge in the matching. See figure 3 for an illustration.

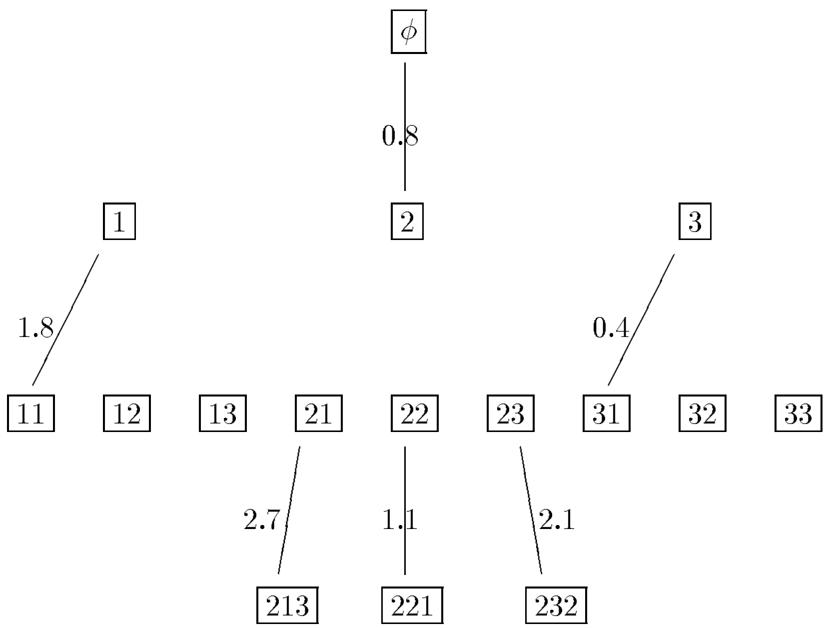

Figure 3. Part of a matching $\mathbf{m}$ on the PWIT. Only the edges of $\mathbf{m}$ are drawn.

We study random matchings $\mathcal{M}$ on the PWIT. Such a random matching will be dependent on the edge-weights, so formally we are dealing with a probability measure μ on $\mathbf{W} \times \mathbf{M}$ (the joint distribution of edge-weights and indicators of edges in the matching) with marginal distribution λ on $\mathbf{W}$.

The connection between matchings $\mathcal{M}$ on the PWIT and matchings π_n in the $n \times n$ random assignment problem is via *local convergence* of matchings on $\mathbf{T}$ induced by the *unfolding map* described informally in section 3; these ideas are explained precisely at the end of this section. Theorem 11 below says (roughly speaking) that $\lim_n EA_n$ is the average cost per edge in a minimum-cost matching on $\mathbf{T}$. At first sight such a result looks false, because there is a greedy matching $\mathcal{M}_{\text{greedy}}$ consisting of edges $(\phi, 1)$, $(2, 21)$, $(3, 31), \ldots, (11, 111)$, $(12, 121), \ldots\ldots$, and $\mathcal{M}_{\text{greedy}}$ has average cost equal to 1, which is certainly not the desired $\lim_n EA_n$. But the precise result is more complicated, because matchings in Theorem 11 are required to satisfy a certain property: informally

(*) The rule for whether an edge e is in the matching should be spatially invariant, that is should not depend on which vertex was chosen as the root of $\mathbf{T}$.

Intuitively, $\mathcal{M}_{\text{greedy}}$ does not satisfy (*), and we verify in section 5.1 that it does not satisfy the precise definition below.

We now work towards making the notion of spatial invariance precise. Our definitions are superficially different from those in [2], but we reconcile them in section 5.5.

For each $\mathbf{w} \in \mathbf{W}$ and each $i \geq 1$ we define an isomorphism of $\mathbf{T}$ (that is, a bijection $\mathbf{V} \to \mathbf{V}$ which induces a bijection $\mathbf{E} \to \mathbf{E}$; we denote either bijection by $\theta_i^{\mathbf{w}}(\cdot)$) which expresses the idea "make vertex i the root, and relabel vertices to preserve order structure". Precisely, first define $k \in \{1, 2, 3, \ldots\}$ by

$$w(i, i(k-1)) < w(\phi, i) < w(i, ik)$$

interpreting the left as 0 for $k = 1$. Then define

$\theta_i^{\mathbf{w}}(i) = \phi$

$\theta_i^{\mathbf{w}}(\phi) = k$

$\theta_i^{\mathbf{w}}$ takes vertices $(i1, i2, \ldots, i(k-1); ik, i(k+1), \ldots)$ to vertices $(1, 2, \ldots, k-1; k+1, k+2, \ldots)$

for $v = ij_2 \ldots j_l$ and $l \geq 3$ let $\theta_i^{\mathbf{w}}(v) = (\theta_i^{\mathbf{w}}(ij_2))j_3 \ldots j_l$

for $v = j_1 j_2 \ldots j_l$ where $j_1 \neq i$ let $\theta_i^{\mathbf{w}}(v) = kj_1j_2 \ldots j_l$.

As already mentioned, the space $\mathbf{W} \times \mathbf{M}$ describes possible combinations $(\mathbf{w}, \mathbf{m})$ of edge-weights $\mathbf{w}$ and a matching $\mathbf{m}$. Consider the enlarged state space $\mathbf{Z} = \mathbf{W} \times \mathbf{M} \times \{1, 2, 3, \ldots\}$ with elements $(\mathbf{w}, \mathbf{m}, k)$ where we interpret the k as meaning that vertex k is distinguished. The maps $\theta_i^{\mathbf{w}}$ induce a single map $\theta : \mathbf{Z} \to \mathbf{Z}$ which we interpret as "relabel the distinguished vertex as the root, relabel other vertices to preserve order structure, and make the previous root into the distinguished vertex". Precisely, $\theta(\mathbf{w}, \mathbf{m}, i) = (\hat{\mathbf{w}}, \hat{\mathbf{m}}, k)$ where

$w(e) = \hat{w}(\theta_i^{\mathbf{w}}(e))$

$m(e) = \hat{m}(\theta_i^{\mathbf{w}}(e))$

$k = \theta_i^{\mathbf{w}}(\phi)$.

See figure 4. Note $\theta = \theta^{-1}$.

A probability measure μ on $\mathbf{W} \times \mathbf{M}$ (describing a random matching on the PWIT) extends to a σ-finite measure $\mu^* := \mu \times \texttt{count}$ on $\mathbf{Z}$, where $\texttt{count}$ is counting measure on $\{1, 2, 3, \ldots\}$.

Definition 10 *A random matching $\mathcal{M}$ on the PWIT, with distribution μ (say) on $\mathbf{W} \times \mathbf{M}$, is* spatially invariant *if μ^* is invariant under θ, that is if $\mu^*(\cdot) = \mu^*(\theta^{-1}(\cdot))$ as σ-finite measures on $\mathbf{Z}$.*

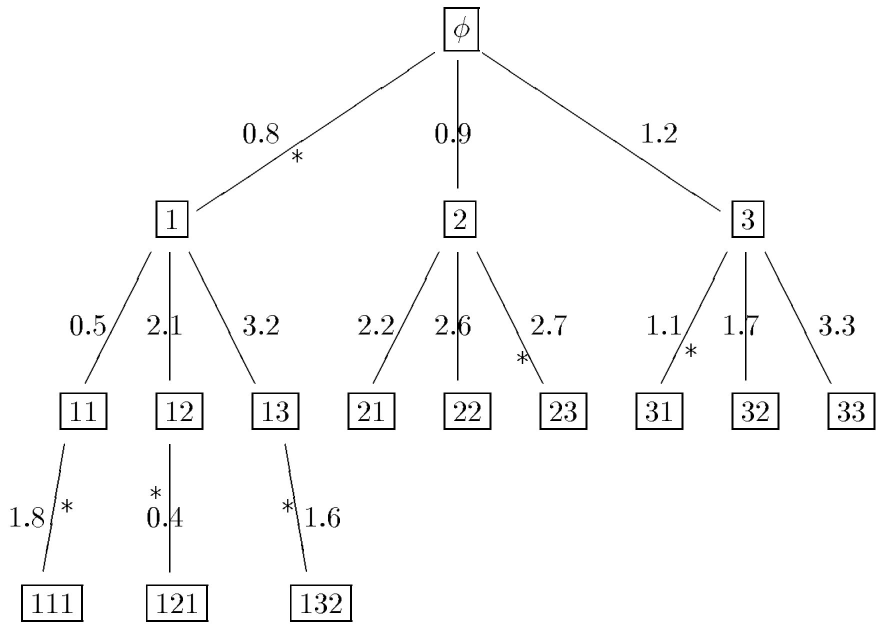

Figure 4. Take the edge-weights $\mathbf{w}$ in figure 2, and the matching $\mathbf{m}$ in figure 3, and distinguish vertex 2. Then figure 4 illustrates part of $\theta(\mathbf{w}, \mathbf{m}, 2) = (\hat{\mathbf{w}}, \hat{\mathbf{m}}, 1)$. Figure 4 shows the new edge-weights $\hat{w}$ attached to the edges; edges in the new matching $\hat{\mathbf{m}}$ are indicated by $*$; the new distinguished vertex is vertex 1.

In a random matching $\mathcal{M}$ the root ϕ is matched to some random neighbor vertex $\overrightarrow{\mathcal{M}}(\phi) \in \{1, 2, 3, \ldots\}$, and the edge $(\phi, \overrightarrow{\mathcal{M}}(\phi))$ has some random cost $W(\phi, \overrightarrow{\mathcal{M}}(\phi))$. We can now reformulate the main result of [2] as

Theorem 11

$$\lim_n EA_n = \inf\{EW(\phi, \overrightarrow{\mathcal{M}}(\phi))\}$$

where the infimum is over all spatially invariant random matchings $\mathcal{M}$ on the PWIT.

Though Theorem 11 is the only result from [2] needed to prove Theorem 1, the other new theorems need further infrastructure from [2]. The reader should perhaps skip the rest of this section on first reading.

A random matching in the $n \times n$ random assignment problem (in brief, an n-matching) was denoted by π_n in the introduction, but can be reformulated as a $n \times n$ $\{0,1\}$-valued random matrix $\mathcal{M} = (m(i,j))$. Call an n-matching *spatially invariant* if the joint distribution $((c(i,j), m(i,j)), 1 \le i,j \le n)$ is invariant under the automorphisms of the complete bipartite graph, that is under permutations of i, under permutations of j, and under complete interchange of i and j. Given any random n-matching, by applying a uniform random automorphism one obtains a spatially invariant n-matching with the same distribution of average-cost-per-edge A_n, so there is no loss of generality in considering only spatially invariant n-matchings. There is a notion of "local convergence" $\mathcal{M}_n \to^d_{\text{local}} \mathcal{M}$, made precise at (13) later, which implies in particular that $c(1, \vec{\mathcal{M}}_n(1)) \stackrel{d}{\to} W(\phi, \vec{\mathcal{M}}(\phi))$. The following two results (stated slightly differently in [2] – see section 5.5) immediately imply Theorem 11.

Theorem 12 *Let $\mathcal{M}$ be a spatially invariant matching on the PWIT with $EW(\phi, \vec{\mathcal{M}}(\phi)) < \infty$. Then there exist spatially invariant n-matchings $\mathcal{M}_n$ such that $\mathcal{M}_n \to^d_{\text{local}} \mathcal{M}$ and $(c(1, \vec{\mathcal{M}}_n(1)),\ 1 \le n < \infty)$ is uniformly intergrable.*

Theorem 13 *Let $n_j \uparrow \infty$, and let $\mathcal{M}_{n_j}$ be spatially invariant n_j-matchings such that* $\limsup_j Ec(1, \vec{\mathcal{M}}_{n_j}(1)) < \infty$. *Then there exists a subsequence n_k of n_j such that $\mathcal{M}_{n_k} \to^d_{\text{local}} \mathcal{M}$, where $\mathcal{M}$ is some spatially invariant matching on the PWIT.*

Remarks. The proof of Theorem 13 is fairly simple, using compactness arguments. The proof of Theorem 12 in [2] is difficult and lengthy. In light of the new results of this paper, we only need Theorem 12 for $\mathcal{M} = \mathcal{M}_{\text{opt}}$, and perhaps the explicit structure of $\mathcal{M}_{\text{opt}}$ could be used to simplify the proof of Theorem 12.

The formalization of *local convergence* involves the *unfolding map* described informally in section 3, which we now describe formally. To formalize *local convergence* requires further notation. Fix n and let $\mathbf{V}^{(n)}$ be the set of vertices $v = v_1 v_2 \ldots v_l$ with $v_1 \le n$ and $v_i \le n-1$ for $i \ge 2$. Write $\mathbf{T}^{(n)} = (\mathbf{V}^{(n)}, \mathbf{E}^{(n)})$ for the corresponding subtree of $\mathbf{T}$. Let G_{nn} be the complete bipartite graph, with vertex-set $\{1,2,\ldots,n\} \times \{1,2\}$. Given a realization $\mathbf{c} = (c(i,j))$ of the cost matrix, regarded as edge-weights on G_{nn}, one can define a graph homomorphism $\psi = \psi_{\mathbf{c}}$ from $\mathbf{T}^{(n)}$ onto G_{nn} as follows.

$\psi(\phi) = (1,1)$

for $i \in \{1,2,3,\ldots,n\}$, define $\psi(i) = (j,2)$, for the j such that $c(1,j)$ is the i'th smallest of $\{c(1,u), 1 \le u \le n\}$

for $i \in \{1,2,3,\ldots,n-1\}$ and $i' \in \{1,2,\ldots,n\}$, define $\psi(i'i) = (k,1)$, for the k such that $c(k,1)$ is the i'th smallest of $\{c(u,\psi(i')), 1 \leq u \leq n, u \neq 1\}$

and then inductively: for $v = v_1 v_2 \ldots v_{2m}$, for $i \in \{1,2,3,\ldots,n-1\}$, define $\psi(vi) = (j,2)$, for the j such that $c(\psi(v),j)$ is the i'th smallest of $\{c(\psi(v),u), 1 \leq u \leq n,\ (u,2) \neq \psi(v^-)\}$ where v^- is the parent of v

for $v = v_1 v_2 \ldots v_{2m+1}$, for $i \in \{1,2,3,\ldots,n-1\}$, define $\psi(vi) = (k,1)$, for the k such that $c(k,\psi(v))$ is the i'th smallest of $\{c(u,\psi(v)), 1 \leq u \leq n,\ (u,1) \neq \psi(v^-)\}$.

This *folding* map $\psi_{\mathbf{c}}$ induces an *unfolding* map which uses the matrix $\mathbf{c}$ to define edge-weights $(W^{(n)}(e))$ on the edge-set $\mathbf{E}^{(n)}$:

$$W^{(n)}(v,vk) = c(i,j) \qquad \begin{array}{l} \text{if } \psi_{\mathbf{c}}(v) = (i,1) \text{ and } \psi_{\mathbf{c}}(vk) = (j,2) \\ \text{or if } \psi_{\mathbf{c}}(v) = (j,2) \text{ and } \psi_{\mathbf{c}}(vk) = (i,1). \end{array}$$

Recall figure 1.

Now for $h \geq 1$ define $\mathbf{E}_{(h)} \subset \mathbf{E}$ as the set of edges each of whose end-vertices is of the form $v_1 v_2 \ldots v_l$ with $l \leq h$ and $\max_i v_i \leq h$. So $\mathbf{E}_{(h)}$ is a finite set of edges. It is straightforward to see that for fixed h

$$(W^{(n)}(e), e \in \mathbf{E}_{(h)}) \ \stackrel{d}{\to}\ (W(e), e \in \mathbf{E}_{(h)})$$

where the limits are the edge-weights in the PWIT. (Essentially, this is the fact that the order statistics of n independent exponential (mean n) r.v.'s converge to the points of a Poisson (rate 1) process.) Recall that we represent a random n-matching as $\{0,1\}$-valued random variables $\mathcal{M}_n(e)$ indexed by the edges e of G_{nn}. Use the homomorphism ψ (considered as a map on edges) to define

$$\widetilde{\mathcal{M}}_n(e) = \mathcal{M}_n(\psi(e)),\ e \in \mathbf{E}^{(n)}.$$

We can now define *local convergence* $\mathcal{M}_n \ \to^d_{\text{local}}\ \mathcal{M}$ to mean: for each fixed h,

$$\left((W^{(n)}(e), \widetilde{\mathcal{M}}_n(e)),\ e \in \mathbf{E}_{(h)}\right) \ \stackrel{d}{\to}\ \left((W(e), \mathcal{M}(e)),\ e \in \mathbf{E}_{(h)}\right). \qquad (13)$$

4.2 Heuristics for the construction of $\mathcal{M}_{\text{opt}}$

Underlying the rigorous construction in the next section is a simple heuristic idea. Given a realization of the PWIT, consider defining

$$\begin{aligned} X_\phi \ &= \ \text{cost of optimal matching on } \mathbf{T} \\ &- \ \text{cost of optimal matching on } \mathbf{T} \setminus \{\phi\}. \end{aligned} \qquad (14)$$

Here we mean *total* cost, so we get $\infty - \infty$, and so (14) makes no sense as a rigorous definition, though statistical physics uses such renormalization arguments all the time. But pretend (14) does make sense. Then for each $v \in \mathbf{V}$ we can define X_v similarly in terms of the subtree $\mathbf{T}^v$ rooted at v (the vertices of $\mathbf{T}^v$ are v and its descendants):

$$\begin{aligned} X_v &= \text{cost of optimal matching on } \mathbf{T}^v \\ &- \text{cost of optimal matching on } \mathbf{T}^v \setminus \{v\}. \end{aligned}$$

And we get the recursion

$$X_v = \min_{1 \leq j < \infty} (W(v, vj) - X_{vj}) \tag{15}$$

because the left side is the cost difference between using or not using v in a matching on $\mathbf{T}^v$; to use an edge (v, vj) we have to pay the cost of the edge and the difference between the cost of not using or using vj, which is the right side. Moreover in the optimal matching on $\mathbf{T}^v$, vertex v should be matched to the vertex vj attaining the minimum in (15).

In the next section we show that one can make a rigorous argument by first constructing (by fiat) random variables satisfying the recursion (15), then using these random variables to define a matching. Not having interpretation (14) means it's not obvious rigorously that this matching is optimal, but it turns out (see start of proof of Proposition 18) that *weak* optimality is quite easy to prove.

4.3 The construction

Each edge $e \in \mathbf{E}$ of $\mathbf{T}$ corresponds to two directed edges $\overrightarrow{e}, \overleftarrow{e}$: write $\overrightarrow{\mathbf{E}}$ for the set of directed edges. For directed edges we have the language of family relationships: each edge $\overrightarrow{e} = (v', v)$ has an infinite number of children of the form (v, y), $y \neq v'$). Thus the directed edge $(273, 27)$ has children $(27, 2)$, $(27, 271)$, $(27, 272)$, $(27, 274), \ldots$, but the reversed edge $(27, 273)$ is not a child. The PWIT has edge-weights $W(e)$ on undirected edges, and we write $W(\overrightarrow{e}) = W(\overleftarrow{e}) = W(e)$ for the directed edges $\overrightarrow{e}, \overleftarrow{e}$ corresponding to e.

Lemma 14 *Jointly with the edge-weights* $(W(e), e \in \mathbf{E})$ *of the PWIT we can construct* $\{X(\overrightarrow{e}), \overrightarrow{e} \in \overrightarrow{\mathbf{E}}\}$ *such that*
(i) each $X(\overrightarrow{e})$ *has the logistic distribution*
(ii) for each $\overrightarrow{e}$*, with children* $\overrightarrow{e_1}, \overrightarrow{e_2}, \ldots$ *say,*

$$X(\overrightarrow{e}) = \min_{1 \leq j < \infty} (W(\overrightarrow{e_j}) - X(\overrightarrow{e_j})) \tag{16}$$

Proof. For a vertex $v = i_1 i_2 \dots i_h$ write $|v| = h$. For $h \geq 0$ write

$$\begin{aligned} \overrightarrow{\mathbf{E}}_h &= \{\overrightarrow{e} = (v, vj) : \ |v| = h, j \geq 1\} \\ \overrightarrow{\mathbf{E}}_{\leq h} &= \{\overrightarrow{e} = (v, y) : \ |v| \leq h, |y| \leq h\}. \end{aligned}$$

Take independent logistic random variables $\{X(\overrightarrow{e}) : \overrightarrow{e} \in \overrightarrow{\mathbf{E}}_h\}$, independent of the family $(W(e), e \in \mathbf{E}, \overrightarrow{e} \in \overrightarrow{\mathbf{E}}_{\leq h})$. We can use (16) recursively to define $X(\overrightarrow{e})$ for each $\overrightarrow{e} \in \overrightarrow{\mathbf{E}}_{\leq h}$. In detail, first use (16) recursively with $k = h-1, h-2, \dots, 0$ to define $X(\overrightarrow{e})$ for each $\overrightarrow{e} \in \cup_{k \leq h} \overrightarrow{\mathbf{E}}_k$ and observe that, for each $k \in [1, h]$, the random variables $\{X(\overrightarrow{e}) : \overrightarrow{e} \in \cup_{r=k}^{h} \overrightarrow{\mathbf{E}}_r\}$ are independent of the family $(W(e), e \in \mathbf{E}, \overrightarrow{e} \in \overrightarrow{\mathbf{E}}_{\leq k})$. Then set

$$\overleftarrow{\mathbf{E}}_h = \{\overrightarrow{e} = (vj, v) : \ |v| = h, j \geq 1\}$$

and use (16) recursively with $k = 0, 1, \dots, h-1$ to define $X(\overrightarrow{e})$ for $\overrightarrow{e} \in \overleftarrow{\mathbf{E}}_k$. Lemma 5 and the natural independence structure ensures that each $X(\overrightarrow{e})$ has logistic distribution. The construction specifies a joint distribution for $\{W(e), \overrightarrow{e} \in \overrightarrow{\mathbf{E}}_{\leq h}; \ X(\overrightarrow{e}), \overrightarrow{e} \in \overrightarrow{\mathbf{E}}_h \cup \overrightarrow{\mathbf{E}}_{\leq h}\}$. The Kolmogorov consistency theorem completes the proof. □

Remarks. (a) By modifying on a null set, we shall assume that the minimum in (16) is attained by a unique j.

(b) For an edge $\overrightarrow{e} = (v', v)$ directed away from ϕ, the $X(\overrightarrow{e})$ constructed above formalizes the notion of X_v in the previous section.

(c) The same type of construction can be associated with general fixed-point identities for distributions which have an appropriate format. See e.g. Aldous [3] for its use in studying a model of "frozen percolation" on the infinite binary tree.

The corollary below (whose routine proof we omit) formalizes the independence structure implicit in the construction. For a directed edge $\overrightarrow{e}$, consider the set consisting of all its descendant edges, but not $\overrightarrow{e}$ itself; then write $D(\overrightarrow{e})$ for this set of edges, considered as *undirected* edges.

Corollary 15 *For each $i \geq 1$ let $f_{i,1}, f_{i,2}, \dots \in \mathbf{E}$ and let $\overrightarrow{e_{i,1}}, \overrightarrow{e_{i,2}}, \dots \in \overrightarrow{\mathbf{E}}$. Write $D_i = \cup_j D(\overrightarrow{e_{i,j}}) \cup \{f_{i,1}, f_{i,2}, \dots\}$. If the sets D_i are disjoint as i varies then the σ-fields $\sigma(X(\overrightarrow{e_{i,j}}), W(f_{i,j}), j \geq 1)$ are independent as i varies.*

Remark. (d) It is natural to ask whether $X(\overrightarrow{e})$ is $\sigma(W(e), e \in D(\overrightarrow{e}))$-measurable, in other words whether the influence of the realization of the "boundary" $\overrightarrow{\mathbf{E}}_h$ in the Lemma 14 construction vanishes in the $h \to \infty$ limit. In the terminology of statistical physics, this asks about extremality of free boundary Gibbs states. We have not studied this question carefully, but simulations suggest the answer is positive.

For $v, v' \in \mathbf{V}$ write $v \sim v'$ if (v, v') is an undirected edge. Fix a realization of $(W(e), e \in \mathbf{E};\ X(\vec{e}), \vec{e} \in \vec{\mathbf{E}})$. For each $v \in \mathbf{V}$ define

$$v^* = \arg\min_{v' \sim v}(W(v, v') - X(v, v')) \tag{17}$$

that is, v^* is the vertex attaining the minimum.

Lemma 16 *The set of undirected edges $\{(v, v^*) :\ v \in \mathbf{V}\}$ is a matching on the PWIT.*

Proof. It suffices to check that $(v^*)^* = v$ for each $v \in \mathbf{V}$. Fix v. Then

$$\begin{aligned} W(v, v^*) - X(v, v^*) &< \min_{y \sim v, y \neq v^*}(W(v, y) - X(v, y)) \text{ by definition of } v^* \\ &= X(v^*, v) \text{ by (16).} \end{aligned}$$

In other words

$$X(v, v^*) + X(v^*, v) > W(v, v^*). \tag{18}$$

Suppose $(v^*)^* = z \neq v$; then

$$W(v^*, z) - X(v^*, z) < W(v^*, v) - X(v^*, v).$$

But by (16)

$$X(v, v^*) \leq W(v^*, z) - X(v^*, z).$$

Adding these last two inequalities gives $X(v, v^*) < W(v^*, v) - X(v^*, v)$ which contradicts (18) and establishes the lemma. □

Write $\mathcal{M}_{\mathrm{opt}}$ for the random matching specified by Lemma 16. Note that the argument leading to (18) can be reversed, to give a more symmetric criterion for whether an edge is in $\mathcal{M}_{\mathrm{opt}}$:

$$e \text{ is an edge of } \mathcal{M}_{\mathrm{opt}} \text{ iff } W(e) < X(\vec{e}) + X(\overleftarrow{e}) \tag{19}$$

where $\vec{e}, \overleftarrow{e}$ are the directed edges corresponding to e. It seems intuitively clear that $\mathcal{M}_{\mathrm{opt}}$ should be spatially invariant; we verify this in section 5.1 as Lemma 24.

Remarks. (e) It is not obvious whether $\mathcal{M}_{\mathrm{opt}}$ is a function of the weights $(W(e), e \in \mathbf{E})$ only, or involves additional external randomization. In fact by (19) this question is equivalent to the question in remark (d) above. It is the possible need for external randomization that motivates formalization of "matchings" as probability distributions on $\mathbf{W} \times \mathbf{M}$ rather than as functions $\mathbf{W} \to \mathbf{M}$.

(f) Intuitively, the quantities $X(\vec{e})$ play a role analogous to *dual variables* in linear programming. But we are unable to make this idea more precise.

4.4 Analysis of $\mathcal{M}_{\rm opt}$

Consider the random cost $W(\phi, \vec{\mathcal{M}}_{\rm opt}(\phi))$ of the edge $(\phi, \vec{\mathcal{M}}_{\rm opt}(\phi))$ containing the root ϕ in $\mathcal{M}_{\rm opt}$.

Proposition 17 *(a) The random variable $W(\phi, \vec{\mathcal{M}}_{\rm opt}(\phi))$ has the probability density function $h(\cdot)$ described in Lemma 6, and so $EW(\phi, \vec{\mathcal{M}}_{\rm opt}(\phi)) = \pi^2/6$.*
(b)

$$P(\vec{\mathcal{M}}_{\rm opt}(\phi) = k) = 2^{-k}, \ k \geq 1.$$

Proof. The edge-weights $(W(\phi, i), i \geq 1)$ and the X-values $(X(\phi, i), i \geq 1)$ are distributed as the Poisson process (ξ_i) and the i.i.d. logistics (X_i) in Lemma 5. Using the latter notation and the definition (17) of $\mathcal{M}_{\rm opt}$,

$$W(\phi, \vec{\mathcal{M}}_{\rm opt}(\phi)) = \xi_I, \text{ where } I = \arg\min_{i \geq 1}(\xi_i - X_i).$$

As a sophisticated way to obtain this distribution (there are alternate, elementary ways) fix $0 < y < \infty$ and condition on the event $A_y := \{\exists J : \xi_J = y\}$. Conditionally, the other points $(\xi_i, i \neq J)$ and other X-values $(X_i, i \neq j)$ are distributed as a Poisson process $(\xi'_j, j \geq 1)$ say and i.i.d. logistics $(X'_j, j \geq 1)$ say, independent of X_J, whose conditional distribution remains logistic. So

$$P(I = J | A_y) = P(y - X_J < X') \text{ where } X' = \min_{j \geq 1}(\xi'_j - X'_j).$$

But by Lemma 5 X' has logistic distribution; since it is independent of X_J, we see from the definition of $h(\cdot)$ in Lemma 6 that $P(I = J|A_y) = h(y)$. Then

$$P(\xi_I \in [y, y + dy]) = P(I = J|A_y)P(\text{ some } \xi_i \text{ in } [y, y + dy]) = h(y)dy$$

establishing part (a). For part (b) we need to show

$$P(I \geq k + 1) = 2^{-k}, \ k \geq 1. \tag{20}$$

From the argument above

$$P(I=J, I \geq k+1 | A_y) = P(y - X_J < \min_{j\geq 1}(\xi'_j - X'_j),\ \xi'_k < y) = h_k(y), \text{ say }.$$

So

$$P(\xi_I \in [y, y+dy], I \geq k+1) = h_k(y)dy$$

and then

$$P(I \geq k+1) = \int_0^\infty h_k(y)dy$$

and the integral is evaluated by Lemma 9.

Proposition 18 *Let $\mathcal{M}$ be a spatially invariant matching of the PWIT such that $P(\vec{\mathcal{M}}(\phi) \neq \vec{\mathcal{M}}_{\rm opt}(\phi)) > 0$. Then $EW(\phi, \vec{\mathcal{M}}(\phi)) > EW(\phi, \vec{\mathcal{M}}_{\rm opt}(\phi))$.*

Before embarking upon the proof of Proposition 18, let us show how Theorems 1 - 4 are deduced from Propositions 17 and 18. Indeed, Propositions 17 and 18 imply

$$\inf\{EW(\phi, \vec{\mathcal{M}}(\phi)) : \vec{\mathcal{M}} \text{ is spatially invariant matching on the PWIT}\} = \tfrac{\pi^2}{6}$$

and so Theorem 11 implies Theorem 1. We next claim that the optimal assignments in the n-matching problem, considered as random matchings $\mathcal{M}_n$ as in section 4.1, satisfy

$$\mathcal{M}_n \ \to^d_{\rm local}\ \mathcal{M}_{\rm opt} \text{ and } (c(1, \vec{\mathcal{M}}_n(1)),\ 1 \leq n < \infty) \text{ is uniformly intergrable.} \tag{21}$$

If not, then by Theorem 13 some subsequence converges locally to some $\mathcal{M} \neq \mathcal{M}_{\rm opt}$, so $\limsup_n EA_n \geq EW(\phi, \vec{\mathcal{M}}(\phi)) > \pi^2/6$, the last inequality by Proposition 18, but this contradicts Theorem 1. Now (21) implies Theorems 2 and 3, because the definition of local convergence (13) implies

$$\left(\vec{\mathcal{M}}_n(1), c(1,1), c(1,2), \ldots\right) \ \stackrel{d}{\to}\ \left(\vec{\mathcal{M}}_{\rm opt}(\phi), W(\phi,1), W(\phi,2), \ldots\right)$$

and then Proposition 17 identifies the required limit distributions. To prove Theorem 4, fix $\delta > 0$. Suppose the assertion were false for that δ; then we can find random n_j-matchings $\mathcal{M}'_{n_j}$ (for some $n_j \to \infty$) such that

$$P\left(\vec{\mathcal{M}}'_{n_j}(1) \neq \vec{\mathcal{M}}_{n_j}(1)\right) \geq \delta \tag{22}$$

where as above $\mathcal{M}_{n_j}$ is the optimal j_n-matching; and

$$Ec\left(1, \vec{\mathcal{M}}'_{n_j}(1)\right) \to \pi^2/6. \tag{23}$$

But by Theorem 13 we can, after passing to a further subsequence, assume $\mathcal{M}'_{n_j} \to^d_{\text{local}} \mathcal{M}$, for some spatially invariant random matching $\mathcal{M}$. Then (22) implies $P(\vec{\mathcal{M}}(\phi) \neq \vec{\mathcal{M}}_{\text{opt}}(\phi)) \geq \delta$, while (23) implies $EW(\phi, \vec{\mathcal{M}}(\phi)) \leq EW(\phi, \vec{\mathcal{M}}_{\text{opt}}(\phi))$, contradicting Proposition 18. This "proof by contradiction" establishes Theorem 4.

Examining the argument above, we see that a weak inequality in Proposition 18 would be sufficient to prove Theorem 1, while the strict inequality is needed for our proofs of Theorems 2 - 4.

4.5 Proof of Proposition 18.

Let $\mathcal{M}$ be a spatially invariant matching, and write A for the event $\{\vec{\mathcal{M}}(\phi) \neq \vec{\mathcal{M}}_{\text{opt}}(\phi)\}$. On A write

$$(v_{-1}, v_0, v_1) = (\vec{\mathcal{M}}_{\text{opt}}(\phi), \phi, \vec{\mathcal{M}}(\phi)).$$

Then on A we can define a doubly-infinite alternating path $\ldots, v_{-2}, v_{-1}, v_0, v_1, v_2, \ldots$ by: $\forall -\infty < m < \infty$

$$(v_{2m-1}, v_{2m}) \text{ is an edge of } \mathcal{M}_{\text{opt}}$$

$$(v_{2m}, v_{2m+1}) \text{ is an edge of } \mathcal{M}.$$

In section 5.1 we shall use spatial invariance to prove

Lemma 19 *Conditional on A, the distributions of $X(v_{-2}, v_{-1})$ and $X(v_0, v_1)$ are the same.*

Now

$$\begin{aligned} X(v_{-2}, v_{-1}) &= \min_{y \sim v_{-1}, y \neq v_{-2}} (W(v_{-1}, y) - X(v_{-1}, y)) \text{ by (16)} \\ &= (W(v_{-1}, v_0) - X(v_{-1}, v_0)) \end{aligned} \tag{24}$$

by (17), because (v_{-1}, v_0) is in $\mathcal{M}_{\text{opt}}$. Also, by (16)

$$X(v_{-1}, v_0) \leq W(v_0, v_1) - X(v_0, v_1) \tag{25}$$

so that

$$D := W(v_0, v_1) - X(v_0, v_1) - X(v_{-1}, v_0) \geq 0.$$

Combining (24) with this definition of D and using $W(v_0, v_{-1}) = W(v_{-1}, v_0)$ gives

$$W(v_0, v_1) - W(v_0, v_{-1}) = D + X(v_0, v_1) - X(v_{-2}, v_{-1}).$$

So

$$\begin{aligned} & EW(\phi, \vec{\mathcal{M}}\ (\phi)) - EW(\phi, \vec{\mathcal{M}}_{\mathrm{opt}}(\phi)) \\ = \quad & E(W(v_0, v_1) - W(v_0, v_{-1}))1_A \\ = \quad & ED1_A + EX(v_0, v_{-1})1_A - EX(v_{-2}, v_{-1})1_A \\ = \quad & ED1_A \text{ by Lemma 19.} \end{aligned} \tag{26}$$

Since $D \geq 0$ this is enough to establish the weak inequality $EW(\phi, \vec{\mathcal{M}}\ (\phi)) \geq EW(\phi, \vec{\mathcal{M}}_{\mathrm{opt}}(\phi))$. To establish the strict inequality, suppose that to the contrary $EW(\phi, \vec{\mathcal{M}}\ (\phi)) = EW(\phi, \vec{\mathcal{M}}_{\mathrm{opt}}(\phi))$. Then (26) implies $ED1_A = 0$, and then (25) implies that on A we have $X(v_{-1}, \phi) = W(\phi, v_1) - X(\phi, v_1)$. And this implies that on A we have

$$v_1 = \arg\min_i{}^{[2]}(W(\phi, i) - X(\phi, i))$$

where $\min^{[2]}$ denotes the second-smallest value, because $\vec{\mathcal{M}}_{\mathrm{opt}}(\phi) = v_{-1} = \arg\min_i(W(\phi, i) - X(\phi, i))$. So without restricting to A we have

$$P\Big(\vec{\mathcal{M}}\ (\phi) = \arg\min_i(W(\phi, i) - X(\phi, i)) \text{ or } \arg\min_i{}^{[2]}(W(\phi, i) - X(\phi, i))\Big) = 1.$$

By an immediate use of spatial invariance we must have the same property at every $v \in \mathbf{V}$:

$$P\left(\vec{\mathcal{M}}\ (v) = \arg\min_i(W(v, i) - X(v, i)) \text{ or } \arg\min_i{}^{[2]}(W(v, i) - X(v, i))\right) = 1. \tag{27}$$

So it's enough to show this can't happen.

Proposition 20 *The only spatially invariant random matching on the PWIT satisfying (27) is* $\mathcal{M}_{\mathrm{opt}}$.

Remark. One can regard Proposition 20 as a kind of "subcritical percolation" fact. There is some set $\mathcal{M}_{2-\mathrm{opt}}$ of edges (v', v'') such that

$$v'' = \arg\min_{v \sim v'}{}^{[2]}(W(v', v) - X(v', v)) \text{ and } v' = \arg\min_{v \sim v''}{}^{[2]}(W(v'', v) - X(v'', v)).$$

In contrast to $\mathcal{M}_{\mathrm{opt}}$, the first equality here does not imply the second. One can show that Proposition 20 is equivalent to the assertion that there is no infinite path consisting of alternating edges from $\mathcal{M}_{\mathrm{opt}}$ and $\mathcal{M}_{2-\mathrm{opt}}$.

Proof. Define a path $\phi = w_0, w_{-1}, w_{-2}, w_{-3}, \ldots$ inductively by: for $m = 1, 2, \ldots$

$$\begin{aligned} w_{-2m+1} &= \arg\min_{y \sim w_{-2m+2}} (W(w_{-2m+2}, y) - X(w_{-2m+2}, y)) \\ w_{-2m} &= \arg\min_{y \sim w_{-2m+1}}{}^{[2]}(W(w_{-2m+1}, y) - X(w_{-2m+1}, y)). \end{aligned}$$

Then

$$(w_{-2m+1}, w_{-2m+2}) \text{ is an edge of } \mathcal{M}_{\mathrm{opt}}, \text{ each } m = 1, 2, 3, \ldots.$$

Suppose $\mathcal{M}$ were a spatially invariant matching satisfying (27) such that $A := \{\vec{\mathcal{M}}(\phi) \neq \vec{\mathcal{M}}_{\mathrm{opt}}(\phi)\}$ has $P(A) > 0$. Then on A we have

$$(w_{-2m}, w_{-2m+1}) \text{ is an edge of } \mathcal{M}, \text{ each } m = 1, 2, 3, \ldots.$$

(The construction of (w_m) resembles the previous construction of (v_m), but note that the definitions of (w_m) above and of $B_m, \bar{B}_q, B^*$ below involve only the PWIT and the $(X(\vec{e}))$, not any hypothetical matching $\mathcal{M}$.) So by (27) for $v = w_{-2m}$

$$A \subseteq B_m := \{w_{-2m+1} = \arg\min_{y \sim w_{-2m}}{}^{[2]}(W(w_{-2m}, y) - X(w_{-2m}, y))\}$$

and so

$$A \subseteq B := \cap_{m=1}^{\infty} B_m.$$

Writing $\bar{B}_q := \cap_{m=1}^{q} B_m$ we have

$$\text{if } P(A) > 0 \text{ then } \lim_{q \to \infty} P(\bar{B}_{q+1})/P(\bar{B}_q) = 1. \tag{28}$$

In section 5.1 we shall use spatial invariance to prove

Lemma 21 *$P(\bar{B}_{q+1}) = P(\bar{B}_q \cap B^*)$, where*

$$B^* := \{\phi = \arg\min_{y \sim w_1}{}^{[2]}(W(w_1, y) - X(w_1, y))\}$$

where

$$w_1 := \arg\min_{y \sim \phi}{}^{[2]}(W(\phi, y) - X(\phi, y)).$$

Using Lemma 21,

$$\lim_{q \to \infty} \frac{P(\bar{B}_{q+1})}{P(\bar{B}_q)} = \lim_q \frac{P(\bar{B}_q \cap B^*)}{P(\bar{B}_q)} = \frac{P(B \cap B^*)}{P(B)} = P(B^* | B).$$

So by (28), to prove Proposition 20 and hence Proposition 18 it is enough to prove

Lemma 22 *If $P(B) > 0$ then $P(B^*|B) < 1$.*

Remark. Intuitively, Lemma 22 seems clear for the following reason. Event B depends only on what happens on the branch from ϕ through w_{-1}, while B^* depends only on what happens on the branch through w_1. Even though there is dependence between the branches, the dependence shouldn't be strong enough to make B^* conditionally *certain* to happen. Formalizing this idea requires study of the effect of conditioning the PWIT, and we defer completing the proof until section 5.3, though the next lemma is one ingredient of the proof.

Lemma 23 *Define*

$$\begin{aligned} X^{\downarrow} &= \min_{i \geq 1}(W(\phi, i) - X(\phi, i)) \\ I &= \arg\min_{i \geq 1}(W(\phi, i) - X(\phi, i)). \end{aligned}$$

For $-\infty < b < a < \infty$ define

$$g(a, b) = P\left(W(\phi, I) - b > \min_{k \geq 1}{}^{[2]}(W(I, Ik) - X(I, Ik)) \,\middle|\, X^{\downarrow} = a\right).$$

Then $g(a, b) > 0$.

Proof. Note that (16) shows

$$X(\phi, I) = \min_{k \geq 1}(W(I, Ik) - X(I, Ik)).$$

Conditionally on $X(\phi, I) = x$, the other values of $\{W(I, Ik) - X(I, Ik),\ k \geq 1\}$ are the points of a certain inhomogeneous Poisson process on (x, ∞), and so

$$P\left(\min_{k\geq 1}{}^{[2]}(W(I,Ik) - X(I,Ik)) \in [y, y+dy] \,\middle|\, X(\phi, I) = x\right) = \beta_x(y)dy \quad (29)$$

for a certain function $\beta_x(\cdot)$ such that $\beta_x(y) > 0$ for all $y > x$. The quantities in (29) are independent of $W(\phi, I)$ and so (29) remains true if we also condition on $W(\phi, I) = a + x$. So

$$\tilde{g}(a, b, x) :=$$
$$P\left(W(\phi, I) - b > \min_{k\geq 1}{}^{[2]}(W(I, Ik) - X(I, Ik)) \,\middle|\, X(\phi, I) = x,\ W(\phi, I) = a + x\right)$$

satisfies $\tilde{g}(a, b, x) > 0$ for all $-\infty < b < a < \infty$ and $-\infty < x < \infty$. Since $X^{\downarrow} = W(\phi, I) - X(\phi, I)$ we see

$$g(a, b) = E(\tilde{g}(a, b, X(\phi, I))|X^{\downarrow} = a) > 0$$

as required.

5 Some technical details

5.1 Spatial invariance proofs

The definition (Definition 10) of spatial invariance for a random matching $\mathcal{M}$ on the PWIT involves the (probability) distribution μ of $((\mathcal{M}(e), W(e)),\ e \in \mathbf{E})$ and the σ-finite measure $\mu^* := \mu \times \texttt{count}$ on $\mathbf{Z}$. Intuitively, if we use an arbitrary rule to distinguish some vertex k (the rule depending on the realization of $\mathcal{M}$ and the $W(e)$), then the μ^*-distribution of the resulting configuration on $\mathbf{Z}$ is just the μ-distribution of the configuration without the vertex being distinguished.

To see the point of spatial invariance, let us show that the greedy matching $\mathcal{M}_{\text{greedy}}$ from section 4.1 is not spatially invariant.

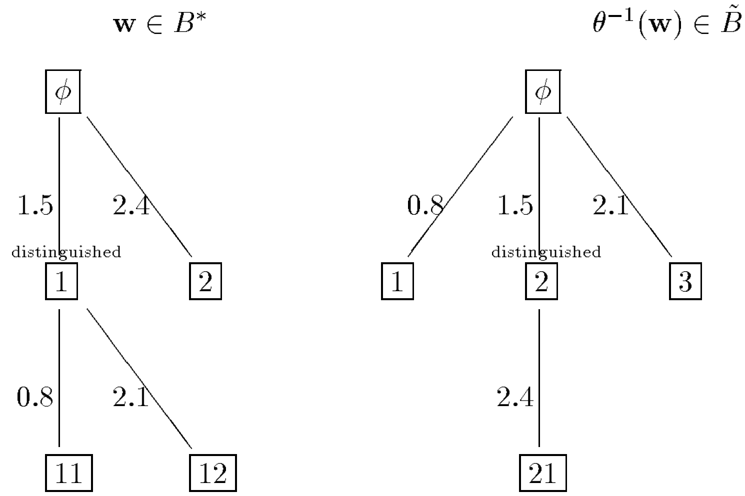

Figure 5. Part of a realization of $\mathbf{w} \in B^*$ and of $\theta^{-1}(\mathbf{w}) \in \tilde{B}$.

Consider the event $B := \{W(1, 11) < W(\phi, 1) < W(1, 12)\}$ which has probability $1/4$. So $B^* := B \cap \{1 \text{ is distinguished }\}$ has μ^*-measure $1/4$. The inverse image $\tilde{B} = \theta^{-1}(B^*)$ is (see figure 5)

$$\tilde{B} = \{W(\phi, 2) < W(2, 21),\ 2 \text{ is distinguished}\}$$

and this has μ^*-measure $1/4$ because $P(W(\phi, 2) < W(2, 21)) = 1/4$. (This equality of μ^*-measures is a consequence of the fact that the distribution of edge-weights is spatially invariant.) Now for a random matching $\mathcal{M}$ to be spatially invariant, the event

$$B^* \cap \{\vec{\mathcal{M}}\ (\phi) = 1\}$$

must have the same μ^*-measure as its inverse image under θ, which is the event

$$\tilde{B} \cap \{\vec{\mathcal{M}}\ (\phi) = 2\}.$$

But for the greedy matching we always have $\vec{\mathcal{M}}\ (\phi) = 1$, so the former event has μ^*-measure $1/4$ while the latter event has μ^*-measure 0.

Proof of Lemma 19. In words, the idea is that the distribution of $X(v_0, v_1)$ as seen from $\phi = v_0$ is the same (by spatial invariance) as the distribution of $X(v_0, v_1)$ as seen from v_2, which (just by relabeling) is the distribution of $X(v_{-2}, v_{-1})$ as seen from $v_0 = \phi$. Precisely:

$$P(A, X(v_0, v_1) \in \cdot) = P\left(\vec{\mathcal{M}}\ (\phi) \neq \vec{\mathcal{M}}_{\rm opt}(\phi), X(\phi, \vec{\mathcal{M}}\ (\phi)) \in \cdot\right)$$

$$= \mu^* \left(\vec{\mathcal{M}}\ (\phi) \neq \vec{\mathcal{M}}_{\rm opt}(\phi), X(\phi, \vec{\mathcal{M}}\ (\phi)) \in \cdot, \vec{\mathcal{M}}\ (\phi) \text{ is distinguished}\right)$$

$$= \mu^* \left(\vec{\mathcal{M}}\ (\phi) \neq \vec{\mathcal{M}}_{\rm opt}(\phi), X(\vec{\mathcal{M}}\ (\phi), \phi) \in \cdot, \vec{\mathcal{M}}\ (\phi) \text{ is distinguished}\right)$$

using spatial invariance to switch the root from ϕ to $\vec{\mathcal{M}}\ (\phi)$, and noting that the event $\{\vec{\mathcal{M}}\ (\phi) \neq \vec{\mathcal{M}}_{\rm opt}(\phi)\}$ is identical to the event $\{\vec{\mathcal{M}}\ (\vec{\mathcal{M}}\ (\phi)) \neq \vec{\mathcal{M}}_{\rm opt}(\vec{\mathcal{M}}\ (\phi))\}$;

$$= \mu^* \left(\vec{\mathcal{M}}\ (\phi) \neq \vec{\mathcal{M}}_{\rm opt}(\phi), X(\vec{\mathcal{M}}\ (\phi), \phi) \in \cdot, \vec{\mathcal{M}}_{\rm opt}(\phi) \text{ is distinguished}\right)$$

because μ^*-measure doesn't depend on the rule for distinguishing a vertex;

$$= \mu^* \left(\vec{\mathcal{M}}\ (\vec{\mathcal{M}}_{\rm opt}(\phi)) \neq \vec{\mathcal{M}}_{\rm opt}(\vec{\mathcal{M}}_{\rm opt}(\phi)), X(\vec{\mathcal{M}}\ (\vec{\mathcal{M}}_{\rm opt}(\phi)), \vec{\mathcal{M}}_{\rm opt}(\phi)) \in \cdot, \vec{\mathcal{M}}_{\rm opt}(\phi) \text{ is distinguished}\right)$$

using spatial invariance to switch the root from ϕ to $\vec{\mathcal{M}}_{\rm opt}(\phi)$;

$$= \mu^* \left(\vec{\mathcal{M}}\ (\phi) \neq \vec{\mathcal{M}}_{\rm opt}(\phi), X(\vec{\mathcal{M}}\ (\vec{\mathcal{M}}_{\rm opt}(\phi)), \vec{\mathcal{M}}_{\rm opt}(\phi)) \in \cdot, \vec{\mathcal{M}}_{\rm opt}(\phi) \text{ is distinguished}\right)$$

because the events are identical;

$$\begin{aligned} &= P\left(\vec{\mathcal{M}}\ (\phi) \neq \vec{\mathcal{M}}_{\rm opt}(\phi), X(\vec{\mathcal{M}}\ (\vec{\mathcal{M}}_{\rm opt}(\phi)), \vec{\mathcal{M}}_{\rm opt}(\phi)) \in \cdot\right) \\ &= P(A, X(v_{-2}, v_{-1}) \in \cdot) \end{aligned}$$

because $v_{-1} = \vec{\mathcal{M}}_{\rm opt}(\phi)$ and $v_{-2} = \vec{\mathcal{M}}\ (v_{-1})$.

Proof of Lemma 21. This is essentially the same argument as above. In words, the idea is that the probability of $\bar{B}_q \cap B^*$ as seen from $\phi = w_0$ is the same (by spatial invariance) as the probability of $\bar{B}_q \cap B^*$ as seen from $w_2 = \arg\min_{w \sim w_1}(W(w_1, w) - X(w_1, w))$, which (just by relabeling) is the probability of $\bar{B}_{q+1}$ as seen from $w_0 = \phi$.

We leave details to the reader.

Lemma 24 $\mathcal{M}_{\mathrm{opt}}$ *is spatially invariant.*

Proof. $\mathcal{M}_{\mathrm{opt}}$ is determined by the $W(e)$ and the $X(\vec{e})$, and the $X(\vec{e})$ satisfy the deterministic relation (16) which is unaffected by relabeling vertices. Moreover the joint distribution of the $X(\vec{e})$ is determined by the fact (Lemma 14) that the r.v.'s $(X(\vec{e}), \vec{e} \in \vec{\mathbf{E}}_h)$ are independent logistic. In words, we need to show that the property "the $(X(\vec{e}), \vec{e} \in \vec{\mathbf{E}}_h)$ are independent logistic" is preserved under θ.

Fix k and write

$$B = \{(X(\vec{e}), \vec{e} \in \vec{\mathbf{E}}_h) \in C, \ k \text{ is distinguished}\}$$

for arbitrary C in the appropriate range space. Write

$$A_l = \{W(l, k-1) < W(\phi, l) < W(l, k)\}.$$

Then (see figure 6)

$$\theta^{-1}(B) = \cup_l \left[A_l \cap \{l \text{ is distinguished } \} \cap \{(X(\vec{e}), \vec{e} \in \vec{\mathbf{E}}_{h,l}) \in C\} \right]$$

where $\vec{\mathbf{E}}_{h,l}$ is the set of edges (v, vj) with $j \geq 1$ and with $v = j_1 j_2 \ldots j_{h-1}$ for $j_1 \neq l$, or with $v = l j_1 j_2 \ldots j_h$. Since $\sum_l P(A_l) = 1$ and the $X(\vec{e})$ under consideration are independent of A_l, verifying $\mu^*(B) = \mu^*(\theta^{-1}(B))$ reduces to showing that $\{X(\vec{e}), \vec{e} \in \vec{\mathbf{E}}_{h,l}\}$ are independent logistic. But this follows from the construction (Lemma 14 and Corollary 15).

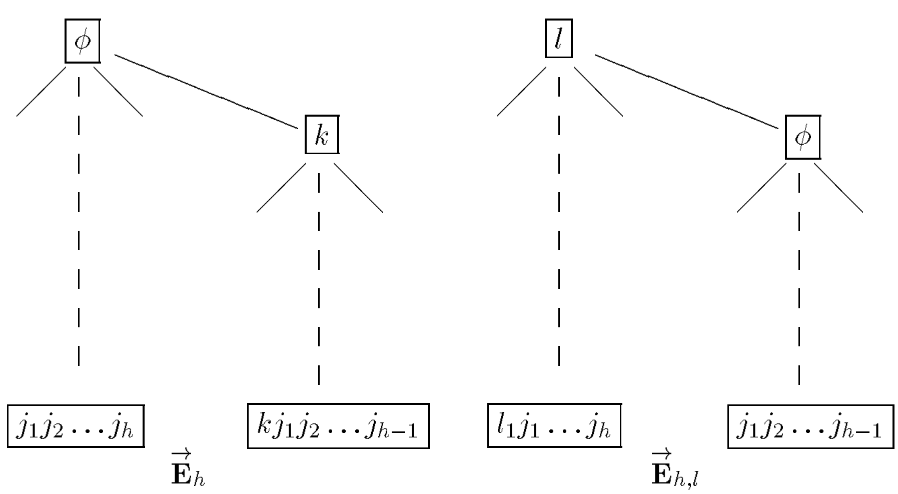

Figure 6.

5.2 The bi-infinite tree

Recall λ denotes the distribution of the edge-weights $(W(e), e \in \mathbf{E})$ of the PWIT. So λ is a probability measure on $\mathbf{W} = (0,\infty)^{\mathbf{E}}$. To incorporate a distinguished neighbor of ϕ we extend to the state space $\mathbf{W} \times \{1,2,3,\ldots\}$ and introduce the σ-finite measure $\lambda \times \mathtt{count}$. We now describe an equivalent way of representing this structure, which turns out to be convenient for certain calculations with $\mathcal{M}_{\mathrm{opt}}$.

Take two copies, $\mathbf{T}^+$ and $\mathbf{T}^-$ say, of the PWIT, and write their vertices as $+v$ and $-v$. Then construct a new "bi-infinite" tree $\mathbf{T}^{\leftrightarrow}$ by joining the roots $+\phi$ and $-\phi$ of $\mathbf{T}^+$ and $\mathbf{T}^-$ via a distinguished edge $(+\phi, -\phi)$. Write $\mathbf{E}^{\leftrightarrow}$ for its edge-set. Let edge-weights on the edges of each of $\mathbf{T}^+$ and $\mathbf{T}^-$ be distributed as in the PWIT, independently for the two sides of $\mathbf{T}^{\leftrightarrow}$. Then define a σ-finite measure $\lambda^{\leftrightarrow}$ on $\mathbf{W}^{\leftrightarrow} := (0,\infty)^{\mathbf{E}^{\leftrightarrow}}$ by specifying that the weight $W(-\phi, +\phi)$ on the distinguished edge should have "distribution" uniform on $(0,\infty)$, independent of the other edge-weights.

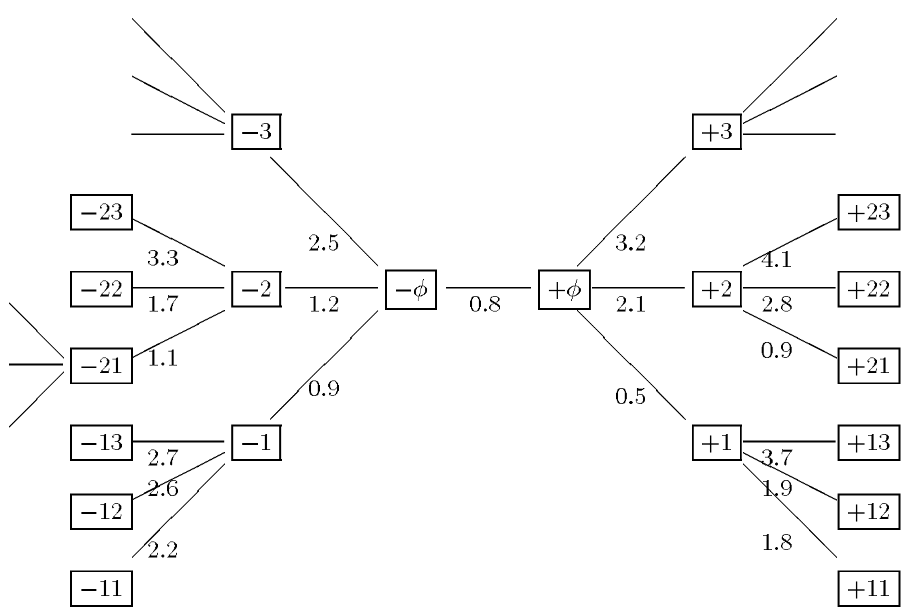

Figure 7. Part of a realization $\mathbf{w}$ of edge-weights on the bi-infinite tree $\mathbf{T}^{\leftrightarrow}$.

The point of this construction is that there is a natural bijection between $\mathbf{W} \times \{1, 2, 3, \ldots\}$ and $\mathbf{W}^{\leftrightarrow}$ which takes $\lambda \times \texttt{count}$ to $\lambda^{\leftrightarrow}$. In fact, if k is the distinguished vertex of $\mathbf{T}$, then relabel vertices according to the rules

relabel k as $-\phi$

relabel j as $+j$ for $j \leq k-1$ and as $+(j-1)$ for $j \geq k+1$

relabel descendants accordingly.

This relabeling induces a map $\psi : \mathbf{W} \times \{1, 2, 3, \ldots\} \to \mathbf{W}^{\leftrightarrow}$ which is invertible. See figure 7. Checking that ψ maps $\lambda \times \texttt{count}$ to $\lambda^{\leftrightarrow}$ reduces to the following easy lemma.

Lemma 25 *Write $\Delta := \{(x_i) : 0 < x_1 < x_2 < \ldots, x_i \to \infty\}$. Write* $\texttt{Pois}$ *for the probability measure on Δ which is the distribution of the Poisson process of rate 1. Consider the map $\chi : \Delta \times \{1, 2, 3, \ldots\} \to \Delta \times (0, \infty)$ which takes $((x_i), k)$ to $((x_i, i \neq k), x_k)$. Then χ maps* $\texttt{Pois} \times \texttt{count}$ *to* $\texttt{Pois} \times \texttt{Leb}$*, where* $\texttt{Leb}$ *is Lebesgue measure on $(0, \infty)$.*

Remark. There is a natural map "reflect" from $\mathbf{W}^{\leftrightarrow}$ to $\mathbf{W}^{\leftrightarrow}$, induced by the bijection of vertices $+v \leftrightarrow -v$. The map $\theta : \mathbf{W} \times \{1, 2, 3, \ldots\} \to \mathbf{W} \times \{1, 2, 3, \ldots, \}$ featuring in the definition of spatial invariance (where we now ignore matchings) is related to ψ as indicated in the diagram.

$$
\begin{array}{ccc}
\mathbf{W} \times \{1, 2, \ldots\} & \stackrel{\psi}{\to} & \mathbf{W}^{\leftrightarrow} \\
\downarrow \theta & & \downarrow \text{reflect} \\
\mathbf{W} \times \{1, 2, \ldots\} & \stackrel{\psi^{-1}}{\leftarrow} & \mathbf{W}^{\leftrightarrow}
\end{array}
$$

In [2], our Theorem 11 was stated in terms of matchings on $\mathbf{T}^{\leftrightarrow}$ instead of $\mathbf{T}$. In this paper we switched to using $\mathbf{T}$ as the basic limit structure for two reasons:

(i) on $\mathbf{T}$ we can define a random matching $\mathcal{M}$ using probability distributions (instead of having to introduce σ-finite measures);

(ii) the definition of spatial invariance (which in either setting involves σ-finite measures) is simpler on $\mathbf{T}$ than on $\mathbf{T}^{\leftrightarrow}$.

In section 5.5 we reconcile the definitions.

The relabeling used to define ψ can also be used to define a family $(X(\overrightarrow{e}), \overrightarrow{e}$ a directed edge of $\mathbf{T}^{\leftrightarrow})$ in terms of the $X(\overrightarrow{e})$ on the directed edges of the PWIT, constructed in Lemma 14. It is easy to check that the joint distribution of $(W(e), X(\overrightarrow{e}); e, \overrightarrow{e}$ edges of $\mathbf{T}^{\leftrightarrow})$ thus obtained is the same as if we applied the construction in Lemma 14 to $(W(e), e \in \mathbf{E}^{\leftrightarrow})$, replacing $\vec{\mathbf{E}}_h$ in the construction by

$$\overset{\leftrightarrow}{bE}_h = \{\vec{e} = (+v, +vj) : \ |v| = h, j \geq 1\} \cup \{\vec{e} = (-v, -vj) : \ |v| = h, j \geq 1\}. \tag{30}$$

Then the matching $\mathcal{M}_{\text{opt}}$ can be defined on $\mathbf{T}^{\leftrightarrow}$ in the same way as on $\mathbf{T}$:

$$e \text{ is an edge of } \mathcal{M}_{\text{opt}} \text{ iff } W(e) < X(\vec{e}) + X(\overleftarrow{e}).$$

Lemma 26 below shows how working on the bi-infinite tree is useful for calculations. Informally, Lemma 26 describes the distribution of $\mathcal{M}_{\text{opt}}$ as seen from a typical edge in $\mathcal{M}_{\text{opt}}$, and exhibits a conditional independence property for the restrictions of $\mathcal{M}_{\text{opt}}$ to the two sides of the tree determined by that edge.

On the PWIT define $X^{\downarrow} = \min_{i \geq 1}(W(\phi, i) - X(\phi, i))$, and write ν_x for the conditional distribution of the family $(W(e), e \in \mathbf{E}; X(\vec{e}), \vec{e} \in \vec{\mathbf{E}}, \vec{e}$ directed away from $\phi)$ given $X^{\downarrow} = x$. Returning to the bi-infinite tree, write λ^1 for the measure obtained by restricting $\lambda^{\leftrightarrow}$ to the set $\{W(-\phi, +\phi) < X(-\phi, +\phi) + X(+\phi, -\phi)\}$. So under λ^1, in $\mathcal{M}_{\text{opt}}$ the vertex $+\phi$ is a.s. matched with vertex $-\phi$.

Lemma 26 *λ^1 is a probability measure. Under λ^1 we have:*
(i) the joint density of $(W(-\phi, +\phi), X(-\phi, +\phi), X(+\phi, -\phi))$ at (w, x_1, x_2) is equal to $f(x_1)f(x_2)1_{(0<w<x_1+x_2)}$, where f is the logistic density;
(ii) conditional on $(W(-\phi, +\phi), X(-\phi, +\phi), X(+\phi, -\phi)) = (w, x_1, x_2)$, the distribution of the family

$$(W(e), e \in \mathbf{E}^+; X(\vec{e}), \vec{e} \in \vec{\mathbf{E}}^+, \vec{e} \ \textit{ directed away from } +\phi)$$

is the image of ν_{x_1} under the natural embedding $\mathbf{T} \to \mathbf{T}^+ \subset \mathbf{T}^{\leftrightarrow}$; the distribution of the family

$$(W(e), e \in \mathbf{E}^-; X(\vec{e}), \vec{e} \in \vec{\mathbf{E}}^-, \vec{e} \ \textit{ directed away from } -\phi)$$

is the image of ν_{x_2} under the natural embedding $\mathbf{T} \to \mathbf{T}^- \subset \mathbf{T}^{\leftrightarrow}$; and these two families are conditionally independent.

Note that because $+\phi$ is matched to $-\phi$, we have $X(+j, +\phi) = W(+\phi, -\phi) - X(+\phi, -\phi)$ and then one can recursively construct $X(\vec{e})$ for $\vec{e}$ directed *toward* $(-\phi, +\phi)$. So the prescription in Lemma 26 is enough to specify the joint distribution of all the $X(\vec{e})$ and hence of $\mathcal{M}_{\text{opt}}$, under λ^1.

Proof of Lemma 26. The joint density has the form stated in (i) by construction; and so its total mass equals $\int\int (x_1+x_2)^+ \ f(x_1)f(x_2)\ dx_1dx_2$ which equals 1 by Lemma 6. Next, by the construction based on $(X(\vec{e}), \vec{e}\in b\overleftrightarrow{E}_h)$ at (30), we see that under $\lambda^{\leftrightarrow}$ the families

$$\{W(e), e\in \mathbf{E}^+; X(\vec{e}), \vec{e}\in\vec{\mathbf{E}}^+, \vec{e} \text{ directed away from } +\phi\} \cup X(-\phi,+\phi)$$

and

$$\{W(e), e\in \mathbf{E}^-; X(\vec{e}), \vec{e}\in\vec{\mathbf{E}}^-, \vec{e} \text{ directed away from } -\phi\} \cup X(+\phi,-\phi)$$

are independent of each other and of $W(-\phi,+\phi)$. Since λ^1 is defined by an event depending only on $\{X(+\phi,-\phi), X(-\phi,+\phi), W(+\phi,-\phi)\}$, we obtain the desired conditional independence property. Each family under $\lambda^{\leftrightarrow}$ is the distributed as the image of the corresponding family on the PWIT (making $X(-\phi,+\phi)$ correspond to $X^{\downarrow}$), and so by independence under $\lambda^{\leftrightarrow}$ the conditional distribution under λ^1 depends only on x_1 (resp. x_2). $\square$

Recall the map $\psi : \mathbf{W}\times\{1,2,3,\ldots\} \to \mathbf{W}^{\leftrightarrow}$ which takes $\lambda\times\texttt{count}$ to $\lambda^{\leftrightarrow}$. The inverse image of the event $\{W(-\phi,+\phi) < X(-\phi,+\phi)+X(+\phi,-\phi)\}$ is

$$\psi^{-1}\{W(-\phi,+\phi) < X(-\phi,+\phi)+X(+\phi,-\phi)\} \ = \ \{\mathcal{M}_{\rm opt}(\phi) \text{ is distinguished}\}.$$

Thus the inverse image of the probability measure λ^1 is $\lambda\times\texttt{count}$ restricted to $\{\mathcal{M}_{\rm opt}(\phi) \text{ is distinguished}\}$. Then after un-distinguishing this vertex, we are left with probability distribution λ on the PWIT.

So in summary, we have established the following "relabeling principle".

> Given $(Z(\vec{e}))$ and hence $\mathcal{M}_{\rm opt}$ on the PWIT, map the whole structure to the bi-infinite tree by relabeling $(\phi, \vec{\mathcal{M}}_{\rm opt}(\phi)$ as $(+\phi, -\phi$ and relabeling other vertices accordingly; then the resulting distribution on the bi-infinite tree is λ^1.

To see why this is useful, let us give a quick second proof of Proposition 17(a). The distribution of $W(\phi, \vec{\mathcal{M}}_{\rm opt}(\phi))$ in the PWIT is the same, by the relabeling principle, as the distribution on the bi-infinite tree of $W(-\phi,+\phi)$ under λ^1. Then by Lemma 26(i)

$$\begin{aligned}\lambda^1\{W(-\phi,+\phi)\in[w,w+dw]\}/dw \ &= \ \int\int_{x_1+x_2>w} f(x_1)f(x_2)\ dx_1dx_2 \\ &= \ P(X_1+X_2>w)\end{aligned}$$

for independent logistic X_1, X_2.

5.3 Calculating with the bi-infinite tree

On the bi-infinite tree define

$$C^* := \{+\phi = \arg\min_{y\sim +I}{}^{[2]}(W(+I,y) - X(+I,y))\}$$

where

$$I = \arg\min_{i\geq 1}(W(+\phi,+i) - X(+\phi,+i)).$$

$$\begin{aligned}
\mathcal{F}^+ &= \sigma(X(\vec{e}), W(e) : e, \vec{e} \text{ edges of } \mathbf{T}^+)\\
\mathcal{F}^- &= \sigma(X(\vec{e}), W(e) : e, \vec{e} \text{ edges of } \mathbf{T}^-)\\
\mathcal{F}^\phi &= \sigma(X(+\phi,-\phi), X(-\phi,+\phi), W(+\phi,-\phi)).
\end{aligned}$$

Lemma 27 *$\lambda^1\{C^{*c}|\mathcal{F}^-,\mathcal{F}^\phi\} = g(X(-\phi,+\phi), W(+\phi,-\phi) - X(+\phi,-\phi))$ for g defined in Lemma 23.*

Proof. C^* is $\mathcal{F}^+$-measurable, so by the conditional independence assertion of Lemma 26 we have

$$\lambda^1\{C^{*c}|\mathcal{F}^-,\mathcal{F}^\phi\} = \lambda^1\{C^{*c}|\mathcal{F}^\phi\}.$$

Thus we have to show that

$$\lambda^1\{C^{*c}|(W(-\phi,+\phi), X(-\phi,+\phi), X(+\phi,-\phi)) = (w,x_1,x_2)\} = g(x_1, w - x_2). \tag{31}$$

Under this conditioning, Lemma 26 implies that the family $(W(e), e \in \mathbf{E}^+; X(\vec{e}), \vec{e} \in \vec{\mathbf{E}}^+, \vec{e}$ directed away from $+\phi)$ is distributed as the image of the family $(W(e), e \in \mathbf{E}; X(\vec{e}), \vec{e} \in \vec{\mathbf{E}}, \vec{e}$ directed away from $\phi)$ conditioned on $\{X^\downarrow = x_1\}$. By definition of $g(a,b)$ in Lemma 23,

$$\begin{aligned}
g(x_1, w-x_2) &= \lambda^1\{W(+\phi,+I) - (w-x_2) > \min_{y\sim +I}{}^{[2]}(W(+I,y) - X(+I,y))\\
&\qquad |(W(-\phi,+\phi), X(-\phi,+\phi), X(+\phi,-\phi)) = (w,x_1,x_2)\}.
\end{aligned}$$

But under this conditioning

$$\begin{aligned}
W(+\phi,+I) - (w-x_2) &= W(+\phi,+I) - (W(+\phi,-\phi) - X(+\phi,-\phi))\\
&= W(+I,+\phi) - X(+I,+\phi)
\end{aligned}$$

by (16), because under λ^1 the vertex $+\phi$ is always matched to $-\phi$. Substituting into (31), we see that the event in (31) is precisely the event C^{*c}, as required.

Proof of Lemma 22. The relabeling principle shows that $P(B^*|B)$ can be rewritten as $\lambda^1\{C^*|C\}$, for a certain event C which is $\mathcal{F}^\phi$-measurable and such that $P(B) = \lambda^1\{C\}$. Now

$$\begin{aligned}
\lambda^1\{C'^{*c}|C\} &= E_{\lambda^1}\, 1_C g(X(-\phi,+\phi), W(+\phi,-\phi) - X(+\phi,-\phi)) \text{ by Lemma 27} \\
&> 0 \text{ if } \lambda^1\{C\} > 0 \text{ by Lemma 23}
\end{aligned}$$

establishing Lemma 22.

5.4 Remarks on quantifying Theorem 4

Recall equation (26): the difference in cost between $\mathcal{M}_{\text{opt}}$ and another spatially invariant random matching $\mathcal{M}$ equals $ED1_A$. To quantify Theorem 4, we would like to show

$$\text{if } P(\vec{\mathcal{M}}(\phi) \neq \vec{\mathcal{M}}_{\text{opt}}(\phi)) = \delta \text{ then } ED1_A \geq \varepsilon(\delta)$$

with some explicit lower bound on $\varepsilon(\delta)$: our current "proof by contradiction" of Theorem 4 shows only that $\varepsilon(\delta) > 0$. Now it is not difficult to improve Lemma 23 to a stronger result giving a lower bound on a corresponding quantity, under the same conditioning. Given some spatially invariant $\mathcal{M}$, one can condition on the one-sided infinite path $+\phi, -\phi, -J, \ldots$ in $\mathbf{T}^{\leftrightarrow}$ whose edges alternate between $\mathcal{M}_{\text{opt}}$ and $\mathcal{M}$, and seek to apply the improved lemma to the conditional distribution of the edge $(+\phi, \vec{\mathcal{M}}(+\phi))$. But the difficulty with this scheme is that the definition of $\mathcal{M}$ might depend on the whole tree; we cannot argue that *a priori* the behavior of $\mathcal{M}$ on $\mathbf{T}^+$ and on $\mathbf{T}^-$ has some conditional independence property, as exploited in the proof of Lemma 27.

5.5 Reconciliation

Here we reconcile the way definitions and results were stated in [2] with the way they are stated in this paper.

Write $\mathbf{W}^* = \mathbf{W} \times \{1, 2, 3, \ldots\}$. The map θ in the definition of spatial invariance is a map $\theta : \mathbf{W}^* \to \mathbf{W}^*$ which preserves $\lambda \times \texttt{count}$. Given a spatially invariant random matching $\mathcal{M}$, define $\gamma : \mathbf{W}^* \to [0,1]$ by

$$\gamma(\mathbf{w}, i) = P(\vec{\mathcal{M}}(\phi) = i | W(e) = w(e)\ \forall e \in \mathbf{E}). \tag{32}$$

Then γ must have the following two properties; the first because $\mathcal{M}$ is a matching, the second because of spatial invariance.

(i) $\sum_i \gamma(\mathbf{w}, i) = 1$.
(ii) $\gamma(\theta(\mathbf{w}, i)) = \gamma(\mathbf{w}, i)$.
And we can write the associated cost as
(iii) $EW(\phi, \vec{\mathcal{M}}(\phi)) = \int_{\mathbf{W}} \sum_i \; w(\phi, i)\gamma(\mathbf{w}, i) \; \lambda(d\mathbf{w})$.

It turns out that we can reverse the argument: given a function $\gamma(\cdot)$ satisfying (i) and (ii), one can define a spatially invariant random matching satisfying (32), whose cost therefore satisfies (iii). Now use the bijection ψ to define

$$g(\mathbf{w}^{\leftrightarrow}) = \gamma(\psi^{-1}(\mathbf{w}^{\leftrightarrow})).$$

Then $g(\cdot)$ satisfies certain consistency conditions corresponding to (i,ii), and the cost (iii) becomes
(iii*) $\quad \int_{\mathbf{W}^{\leftrightarrow}} g(\mathbf{w}^{\leftrightarrow}) \; \lambda^{\leftrightarrow}(d\mathbf{w}^{\leftrightarrow})$.
Thus the quantity $\inf\{EW(\phi, \vec{\mathcal{M}}(\phi)) : \quad \mathcal{M}$ spatially invariant$\}$ in our Theorem 11 equals the infimum of (iii*) over functions g satisfying the consistency conditions, and that was the formulation of the limit constant $\lim_n EA_n$ in [2] Theorem 1.

Our Theorem 13 is [2] Proposition 3(a), stated there for the optimal n-matchings but extending unchanged to arbitrary spatially invariant matchings. Our Theorem 12 is obtained by combining Proposition 3(b) and Proposition 2 of [2]; local convergence is the method of proof of those results.

6 Variants of the random assignment problem

From our asymptotic viewpoint, the non-bipartite matching problem – with $n \times n$ matrix $(c(i, j))$ with $c(i, j) \equiv c(j, i)$ and n even – is the same (up to normalization convention) as the bipartite problem, because the limit random structure and problem is exactly the same matching problem on the PWIT. Below we describe variants where the limit random structure or problem is different. Giving rigorous proofs of existence and uniqueness of solutions of the resulting fixed point distributional equations (39, 42, 46) and rigorous justification of the relation between the finite-n problem and the asserted asymptotic structure, remain open problems.

6.1 Power-law densities

Consider the random assignment problem

$$A_n = \min_{\pi} \sum_{i=1}^{n} c(i, \pi(i))$$

where now the $(c(i,j))$ are i.i.d. with density $f_c(\cdot)$ satisfying

$$f_c(x) \sim x^r \text{ as } x \downarrow 0 \tag{33}$$

for some $0 < r < \infty$. This setting is motivated by trying to "fit" mean-field models to the Euclidean matching problem on random points in $[0,1]^d$, for which the distribution of inter-point distances satisfies (33) with $r = d-1$: see [12] for further discussion. In this setting, EA_n will grow as order $n^{\frac{r}{r+1}}$, and so we rescale the problem to study $A_n' = n^{-\frac{r}{r+1}} A_n$; in other words

$$A_n' = \min_\pi \frac{1}{n} \sum_{i=1}^{n} c'(i, \pi(i))$$

where $c'(i,j) = n^{1/(r+1)} c(i,j)$. Mézard - Parisi ([21] eq. (22,23)) use the replica method to argue $EA_n' \to \gamma_r$ where the limit constant is characterized by

$$\gamma_r = (r+1) \int_{-\infty}^{\infty} G(l) e^{-G(l)} \, dl \tag{34}$$

$$G(l) = \frac{2}{r!} \int_{-l}^{\infty} (l+y)^r e^{-G(y)} \, dy, \quad -\infty < l < \infty. \tag{35}$$

Let us indicate how our approach gives essentially the same result. (The calculations in the rest of section 6.1 derive from notes of Boris Pittel). Underlying the connection between the matrix with i.i.d. exponential (mean n) entries and the PWIT is the fact that the order statistics $(\xi_1^{(n)}, \xi_2^{(n)}, \ldots)$ of n exponential (mean n) r.v.'s satisfy

$$(\xi_1^{(n)}, \xi_2^{(n)}, \ldots) \stackrel{d}{\to} (\xi_1, \xi_2, \ldots) \tag{36}$$

where the limit is a Poisson (rate 1) process. In the current setting, the order statistics $(\xi_i^{(n)})$ of n r.v.'s distributed as $c'(i,j)$ satisfy (36) where the limit (ξ_i) is now the inhomogeneous Poisson process of rate

$$P(\text{ some point of } (\xi_i) \text{ in } [x, x+dx]) = x^r \, dx. \tag{37}$$

Without checking details, it seems clear that following the method of [2] and this paper leads to the formulas which parallel (2,3)

$$\gamma_r = \int_0^\infty x \cdot x^r P(X_1 + X_2 > x) \, dx \tag{38}$$

$$X \stackrel{d}{=} \min_i (\xi_i - X_i) \tag{39}$$

where (ξ_i) is now the inhomogeneous Poisson process (37). The calculations below will check that (34,35) and (38,39) are essentially equivalent. Writing $F(x) = P(X > x)$, (39) says

$$
\begin{aligned}
F(x) &= P(\text{ no point of } (\xi_i, X_i) \text{ lies in } \{(z,b) : \ z - b \leq x\}) \\
&= \exp\left(- \int_0^\infty z^r F(z-x) \ dz \right). \qquad (40)
\end{aligned}
$$

Setting $y = z - x$ gives

$$
F(x) = \exp\left(- \int_{-x}^\infty (y+x)^r F(y) \ dy \right).
$$

Setting $G(x) = -\log F(x)$ gives

$$
G(x) = \int_{-x}^\infty (y+x)^r e^{-G(y)} \ dy
$$

which agrees with (35) up to the constant factor $2/r!$. [why the discrepancy? The 2 is the normalization convention mentioned in section 2. The $r!$ is a typo at [21] eq. (5), where the l^r should be $l^r/r!$ as it is in [22] eq. (2.1).] Next, (38) says

$$
\begin{aligned}
\gamma_r &= \int_0^\infty x^{r+1} dx \int_{-\infty}^\infty P(X_1 \in dy) P(X_2 \geq x - y) \\
&= \int_{-\infty}^\infty P(X \in dy) \int_0^\infty x^{r+1} F(x-y) \ dx \\
&= - \int_{-\infty}^\infty F(y) dy \int_0^\infty x^{r+1} F'(x-y) dx \\
&\quad \text{(integrating by parts over } y) \\
&= \int_{-\infty}^\infty F(y) dy \ (r+1) \int_0^\infty x^r F(x-y) dx \\
&\quad \text{(integrating by parts over } x) \\
&= (r+1) \int_{-\infty}^\infty F(y)(-\log F(y)) \ dy \quad \text{by (40)} \\
&= (r+1) \int_{-\infty}^\infty G(y) e^{-G(y)} \ dy \quad \text{for } G(y) = -\log F(y)
\end{aligned}
$$

and this is exactly (34).

6.2 The combinatorial TSP

In the combinatorial (i.e. symmetric mean-field) traveling salesman problem (TSP) we take a $n \times n$ matrix of "distances" $c(i,j)$ such that $(c(i,j), 1 \leq i < j \leq n)$ are i.i.d. exponential (mean n) and $c(i,i) \equiv 0$ and $c(j,i) \equiv c(i,j)$. We study

$$L_n := \min_\pi \tfrac{1}{n} \sum_i c(i, \pi(i))$$

where π is a cyclic permutation of $\{1, 2, \ldots, n\}$. [An *asymmetric* version of this problem is asymptotically equivalent to the random assignment problem [13], so less interesting from our viewpoint.] One can mimic the heuristic argument of section 4.2. That is, for the subtree $\mathbf{T}^v$ rooted at v, a *tour* is a set of doubly-infinite paths which pass once through each vertex of $\mathbf{T}^v$; an *almost-tour* is the variation in which the root v is the start of one path in the set. Write (heuristically)

$$(*) \qquad \begin{aligned} X_v \;&=\; \text{cost of optimal tour on } \mathbf{T}^v \\ &-\; \text{cost of optimal almost-tour on } \mathbf{T}^v. \end{aligned}$$

One can continue the parallel heuristic to argue that the solution format of the combinatorial TSP is very similar to that for the random assignment problem given by (2,3):

$$\lim_n EL_n = \int_0^\infty zP(X_1 + X_2 \geq z)\, dz \tag{41}$$

where the distribution of X, X_i is determined by

$$X \stackrel{d}{=} \min_i{}^{[2]}(\xi_i - X_i) \text{ where } (\xi_i) \text{ is a Poisson(1) process;} \tag{42}$$

where $\min^{[2]}$ denotes the second-smallest value (so the only difference between the two solutions is the replacement of min by $\min^{[2]}$). By copying the proof of Lemma 5 one can progress toward an explicit solution. Writing $\bar{F}(y) = P(X > y)$, (42) says (in the notation of Lemma 5)

$$\begin{aligned} \bar{F}(y) &= P(\text{ 0 or 1 points of } \mathcal{P} \text{ in } \{(z,x): \; z - x \leq y\}) \\ &= \left(1 + \int_{-y}^\infty \bar{F}(u)du\right) \exp\left(-\int_{-y}^\infty \bar{F}(u)du\right). \end{aligned}$$

Now write $G(y) = \int_{-y}^\infty \bar{F}(u)du$ so that the fixed-point equation (42) becomes the fixed-point equation

$$G'(y) = (1 + G(-y)) \exp(-G(-y)). \tag{43}$$

To rewrite (41) in terms of G, set $z = x_1 + x_2$.

$$\begin{aligned}
\int_0^\infty zP(X_1+X_2>z)\ dz &= \int_0^\infty z\ dz \int_{-\infty}^\infty F'(x_2)\bar{F}(z-x_2)\ dx_2 \\
&= \int_0^\infty z\ dz \int_{-\infty}^\infty F'(z-x_1)\bar{F}(x_1)\ dx_1 \\
&= \int_{-\infty}^\infty \bar{F}(x_1)\ dx_1 \int_0^\infty zF'(z-x_1)\ dz \\
&= \int_{-\infty}^\infty \bar{F}(x_1)\ dx_1 \int_0^\infty F(z-x_1)\ dz \\
&= \int\int_{x_1+x_2>0} \bar{F}(x_1)F(x_2)\ dx_1 dx_2 \\
&= \int_{-\infty}^\infty -G'(x_1)G(-x_1)\ dx_1.
\end{aligned}$$

Now use (43) to see that (41) becomes

$$\lim_n EL_n = \int_{-\infty}^\infty G(x)(1+G(x))e^{-G(x)}\ dx. \qquad (44)$$

So we can rewrite our solution (41,42) as (43,44), which is the solution given by Krauth - Mézard [17] using the cavity method. Earlier work [22] using the replica method gave more complicated formulas. As in section 6.1 one can consider the more general setting (33) and, after some calculus, our solution (with ξ_i now defined via (37)) again coincides with the general-r expression in [17]. Our probabilistic expressions (41,42;2,3) perhaps makes the mathematical similarities between the heuristic solutions of TSP and the random assignment problems more visible than do the analytic expressions in the physics literature. In contrast to the random assignment case, the solution of identity (43) seems to have no simple explicit formula. Numerically solving (43,44) gives a limit constant of about 2.04 [17], in agreement with Monte Carlo simulations of the finite-n combinatorial TSP, and the agreement persists for larger values of r; see Percus - Martin [28] for a recent review.

Returning to the bipartite setting, one can define a k-assignment problem: study

$$A_n^{(k)} := \min_S \tfrac{1}{nk} \sum_{(i,j)\in S} c_{ij}$$

where S denotes an edge-set such that $|\{j : (i,j) \in S\}| = k\ \forall i$ and $|\{i : (i,j) \in S\}| = k\ \forall j$. A very similar heuristic argument indicates that $\lim_n EA_n^{(k)}$ should be given by the analog of (41,42) with k'th minimum in place of second-minimum.

6.3 Gibbs measures on assignments

Fix a parameter $0 < \lambda < \infty$. Take $(c(i,j), 1 \leq i,j \leq n)$ i.i.d. exponential (mean n). Define a non-uniform random permutation Π_n of $\{1,2,\ldots,n\}$ by

$$P(\Pi_n = \pi | \text{ all the } c(i,j)) \propto \exp(-\lambda \sum_i c(i,\pi(i))). \tag{45}$$

The statistical physics approach to the random assignment problem (and other combinatorial optimization problems over random data) is to first study such *Gibbs measures* and then take $\lambda \to \infty$ limits.

It seems plausible that our method extends to the study of the $n \to \infty$, λ fixed, limit behavior of this Gibbs measure, though it may be technically challenging to make these ideas rigorous. In brief, regard the right side of (45) as specifying a random matching Π_∞ on the PWIT, and seek the density $h_\lambda(x)$ of cost-x edges in the matching Π_∞. To mimic the heuristic argument of section 4.2, then regard the right side of (45) as specifying a "matching or almost-matching" Π^+ on $\mathbf{T}^+$, and define

$$X_\phi = \frac{P(\Pi^+ \text{ is an almost-matching })}{P(\Pi^+ \text{ is a matching })}.$$

One can argue heuristically that this X_ϕ should satisfy the distributional identity

$$X \stackrel{d}{=} \left(\sum_{i=1}^{\infty} e^{-\lambda \xi_i} X_i \right)^{-1} \tag{46}$$

for Poisson (rate 1) ξ_i, and that the desired limit density is

$$h_\lambda(x) = 1 - E\left(\frac{1}{1 + X_1 X_2 e^{-\lambda x}} \right). \tag{47}$$

Similar ideas are implicit in Talagrand [33], who obtains some rigorous results for small λ. Mézard - Parisi ([21] eq. (18-20)) derive a non-rigorous solution in a different form; presumably (cf. sections 6.1 and 6.2) the two forms are equivalent but we have been unable to verify this. It is easy to verify that, writing $X = e^{-\lambda U}$ in (46), expressions (46,47) are consistent in the $\lambda \to \infty$ limit with (2,3). Formalizing our approach in the Gibbs setting seems more challenging than in previous settings, for the following reason. A central part of [2] is showing (Proposition 2) that given an almost-complete matching, one can construct a complete matching for small extra cost. Proving an analogous result in the Gibbs setting will presumably require the same type of technical estimates needed in [33].

7 Final remarks

7.1 The AEU property in random optimization

One can define the AEU property in quite general "optimization over random data" settings. Consider for instance the Euclidean TSP involving n i.i.d. points on $[0,1]^2$. Writing $L(T_n)$ for the length of the shortest tour, we have [5] $EL(T_n) \sim cn^{1/2}$ for constant c. The AEU property would be: for non-optimal tours T'_n, if $En^{-1}\#\{e : e \in T_n \setminus T'_n\} \geq \delta$ then $\liminf_n n^{-1/2}(EL(T'_n) - EL(T_n)) \geq \varepsilon(\delta)$ where $\varepsilon(\delta) > 0$. Whether or not the AEU property holds for the Euclidean TSP is a challenging question. Conceptually, the AEU property seems an interesting way of classifying optimization problems from the average-case viewpoint and seems worthy of further study. Note it is an intrinsic property of the problem, not of any particular algorithm to solve the problem.

AEU is an instance of what statistical physicists call *replica symmetry* (RS), the opposite of *replica symmetry breaking* (RSB). Physicists use these terms (as they use *phase transition*) in a variety of contexts; AEU is a formalization of RS in the particular context of optimization.

7.2 The exact conjecture

Our proofs are purely asymptotic, so do not shed light on the exact conjecture (1). Note (cf. proof of Proposition 17) that $h(x)$ represents the asymptotic chance than a cost-x edge gets into the optimal matching, so in particular a cost-0 edge has asymptotic chance $h(0) = 1/2$ to be in the optimal matching, as suggested by previous work ([25], [19] sec. 4.1).

It is very natural to conjecture that var $A_n \sim \sigma^2/n$ for some $0 < \sigma < \infty$, and that rescaled A_n has a Normal limit: this is supported by Monte Carlo simulation [12]. One can define a candidate value $\tilde{\sigma}$ in terms of the optimal matching on the infinite tree, which (roughly speaking) reflects *local* dependence, but it is not apparent even heuristically whether $\tilde{\sigma}$ should coincide with σ.

Parisi [27] observes that $EA_n^2 \approx EA_n + 1$ for small n. The implicit conjecture is an amusing instance of how numerics can mislead; what's really going on is

$$EA_n^2 \approx (EA_n)^2 \text{ for large } n$$

and $\pi^2/6$ just happens to be close to the solution of $z^2 = z + 1$.

7.3 Frieze's $\zeta(3)$ result

It is intriguing that the limit $\zeta(3)$ arises in the somewhat analogous problem of the minimum-weight spanning tree on the complete graph with i.i.d. edge-weights (Frieze [10]). That problem can also be studied using the PWIT (Aldous [1]). The density function $\hat{h}(x)$ with mean $\zeta(3)$ analogous to $h(x)$ is

$$\hat{h}(x) = \tfrac{1}{2}(1 - q^2(x))$$

where $q(x) = 0$ on $0 \leq x \leq 1$ and for $x > 1$ is the non-zero solution of

$$1 - q(x) = \exp(-xq(x)).$$

Thus we see the same structure: calculations of asymptotic quantities involve a fixed-point identity. It is natural to speculate that some other problem has solution $\zeta(4)$.

Acknowledgements. I thank Greg Sorkin, Boris Pittel, Michel Talagrand and Jim Pitman for discussions and encouragement, and a referee for careful comments.

References

[1] D.J. Aldous. A random tree model associated with random graphs. *Random Structures Algorithms*, 1:383–402, 1990.

[2] D.J. Aldous. Asymptotics in the random assignment problem. *Probab. Th. Rel. Fields*, 93:507–534, 1992.

[3] D.J. Aldous. The percolation process on a tree where infinite clusters are frozen. *Math. Proc. Cambridge Philos. Soc.*, 128:465–477, 2000.

[4] S.E. Alm and G. Sorkin. Exact expectations and distributions for the random assignment problem. Technical report, IBM T.J. Watson, 1999.

[5] J. Beardwood, H.J. Halton, and J.M. Hammersley. The shortest path through many points. *Proc. Cambridge Phil. Soc.*, 55:299–327, 1959.

[6] M.W. Buck, C.S. Chan, and D.P. Robbins. On the expected value of the minimum assignment. arXiv: math.CO/0004175, 2000.

[7] D. Coppersmith and G. Sorkin. Constructive bounds and exact expectations for the random assignment problem. *Random Structures Algorithms*, 15:113–144, 1999.

[8] V. S. Dotsenko. Exact solution of the random bipartite matching model. *J. Phys. A*, 33:2015–2030, 2000.

[9] M.E. Dyer, A.M. Frieze, and C.J.H. McDiarmid. On linear programs with random costs. *Math. Programming*, 35:3–16, 1986.

[10] A.M. Frieze. On the value of a random minimum spanning tree problem. *Discrete Appl. Math.*, 10:47–56, 1985.

[11] M.X. Goemans and M.S. Kodialam. A lower bound on the expected cost of an optimal assignment. *Math. Oper. Res.*, 18:267–274, 1993.

[12] J. Houdayer, J.H. Bouter de Monvel, and O.C. Martin. Comparing mean field and Euclidean matching problems. *Eur. Phys. J. (B)*, 6:383–393, 1998.

[13] R.M. Karp. A patching algorithm forthe non-symmetric traveling salesman problem. *SIAM J. Comput.*, 8:561–573, 1979.

[14] R.M. Karp. An upper bound on the expected cost of an optimal assignment. In *Discrete Algorithms and Complexity: Proceedings of the Japan-U.S. Joint Seminar*, pages 1–4. Academic Press, 1987.

[15] R.M. Karp and M. Sipser. Maximum matchings in sparse random graphs. In *Proc. 22nd IEEE Symp. Found. Comp. Sci.*, pages 364–375, 1981.

[16] S. Kotz and N.L. Johnson, editors. *Logistic Distribution*, volume 5 of *Encyclopedia of Statistical Sciences*, pages 123–128. Wiley, 1985.

[17] W. Krauth and M. Mézard. The cavity method and the travelling-salesman problem. *Europhys. Lett.*, 8:213–218, 1987.

[18] A. Lazarus. Certain expected values in the random assignment problem. *Oper. Res. Lett.*, 14:207–214, 1993.

[19] S. Linusson and J. Wästlund. A generalization of the random assignment problem. arXiv: math.CO/0006146, 2000.

[20] R. Lyons and Y. Peres. Probability on trees and networks. Book in preparation, available at http://www.ma.huji.ac.il/ lyons/prbtree.html, 1996.

[21] M. Mézard and G. Parisi. Replicas and optimization. *J. Physique Lett.*, 46:L771–L778, 1985.

[22] M. Mézard and G. Parisi. A replica analysis of the travelling salesman problem. *J. Physique*, 47:1285–1296, 1986.

[23] M. Mézard and G. Parisi. On the solution of the random link matching problem. *J. Physique*, 48:1451–1459, 1987.

[24] M. Mézard, G. Parisi, and M.A. Virasoro. *Spin Glass Theory and Beyond*. World Scientific, Singapore, 1987.

[25] B. Olin. *Asymptotic Properties of Random Assignment Problems*. PhD thesis, Kungl Tekniska Högskolan, Stockholm, Sweden, 1992.

[26] G. Parisi. *Field Theory, Disorder and Simulations*. Number 49 in Lecture Notes in Physics. World Scientific, Singapore, 1992.

[27] G. Parisi. A conjecture on random bipartite matching, 1998. arXiv: cond-mat/9801176.

[28] A.G. Percus and O.C. Martin. The stochastic traveling salesman problem: Finite size scaling and the cavity prediction. *J. Statist. Phys.*, 94:739–758, 1999.

[29] J. Pitman and M. Yor. Infinitely divisible laws associated with hyperbolic functions. Technical Report 581, Dept. Statistics, U.C. Berkeley, 2000.

[30] S.T. Rachev and L. Ruschendorf. Probability metrics and recursive algorithms. *Adv. in Appl. Probab.*, 27:770–799, 1995.

[31] J.M. Steele. *Probability Theory and Combinatorial Optimization.* Number 69 in CBMS-NSF Regional Conference Series in Applied Math. SIAM, 1997.

[32] M. Talagrand. Huge random structures and mean field models for spin glasses. In *Proc. International Congress Mathematicians 1998, Vol. 1 (Berlin, 1998)*, pages 507–536, 1999. Extra vol. of *Documenta Mathematica.*

[33] M. Talagrand. An assignment problem at high temperature. Unpublished, 2000.

[34] M. Talagrand. Mean field models for spin glasses: a first course. To appear in École d'été de Probabilités de Saint-Flour, XXX – 2000, 2000.

[35] D.W. Walkup. On the expected value of a random assignment problem. *SIAM J. Comput.*, 8:440–442, 1979.